جامعة الشرقية
A' SHARQIYAH UNIVERSITY

معادلة درجات امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة في سلطنة عُمان باستخدام نظرية الاستجابة للفقرة

علي بن عامر بن سلّوم الهنائي

رسالة مقدمة استكمالاً لمتطلبات الحصول على درجة
الماجستير في التربية
تخصص: القياس والتقويم

قسم علم النفس
كلية الآداب والعلوم الإنسانية
جامعة الشرقية
سلطنة عُمان

2025م / 1447هـ

جامعة الشرقية
A' SHARQIYAH UNIVERSITY

كلية الآداب والعلوم الإنسانية

قسم علم النفس

معادلة درجات امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة في سلطنة عُمان باستخدام نظرية الاستجابة للفقرة

رسالة مقدمة استكمالًا لمتطلبات الحصول على درجة الماجستير في التربية

تخصص: القياس والتقويم

إعداد

علي بن عامر بن سلّوم الهنائي

لجنة الإشراف

د. شريف بن عبد الرحمن السعودي مشرفًا رئيسًا

د. عامر بن سالم الحبسي مشرفًا ثانيًا

السنة الأكاديمية

2025-2024

قرار لجنة المناقشة

معادلة درجات امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة في سلطنة عُمان باستخدام نظرية الاستجابة للفقرة

أعدها الطالب:

علي بن عامر بن سلّوم الهنائي

نوقشت هذه الرسالة وأجيزت بتاريخ 2025/10/06م

المشرف الرئيس	المشرف الثاني
شريف السعودي	عامر الحبسي

أعضاء لجنة المناقشة

م	صفته في اللجنة	الاسم	الرتبة الأكاديمية	التخصص	الكلية/ المؤسسة	التوقيع
1	رئيس اللجنة	د. عصام اللواتي	أستاذ مشارك	علم النفس التربوي	جامعة الشرقية	
2	المناقش الخارجي	أ.د. حسين الخروصي	أستاذ	القياس والتقويم	جامعة السلطان قابوس	
3	المناقش الداخلي	د. إبراهيم الوهيبي	أستاذ مساعد	القياس والتقويم	جامعة الشرقية	
4	المشرف الرئيس	د. شريف السعودي	أستاذ مشارك	القياس والتقويم	جامعة الشرقية	

الإقـــــــرار:

أُقرّ بأن المادة العلمية الواردة في هذه الرسالة قد حُدِّدَ مصدرها العلميّ، وأن محتوى الرسالة غير مقدم للحصول على أي درجة علمية أخرى، وأن مضمون هذه الرسالة يعكس آراء الباحث الخاصة، وهي ليست بالضرورة الآراء التي تتبناها الجهة المانحة.

الباحث: علي بن عامر بن سلّوم الهنائي الرقم الجامعي: 2318629

التوقيع:

بسم الله الرحمن الرحيم

إهـــــداء

إلى من غرسا في نفسي بذور الطموح، وعلّماني أن الإيمان بالله والعمل الجاد هو طريق النجاح...

إلى والديّ العزيزين، سندي في الحياة، ونور دربي؛ فكلّ الحروف تعجز عن شكر عطائهما وتضحياتهما...

إلى أسرتي الكريمة، منبع الحب والطمأنينة؛ فقد كانت الدافع والداعم في كلّ لحظة...

إلى إخوتي وأخواتي، شركاء الرحلة، ومن كانوا دائمًا الحافز والمشجّع في الأوقات الصعبة...

إلى كلّ من كانت له بصمة في هذا العمل، أستاذًا أو صديقًا أو زميلًا، ممن قدّموا لي يد العون أو كلمة دعم أو رأيٍ سديد...

أهديكم جميعًا هذا العمل المتواضع، تعبيرًا عن شكرٍ وامتنانٍ لا توفّيه الكلمات.

الباحث

شكر وتقدير

الحمد لله أولًا وآخرًا، ظاهرًا وباطنًا، الحمد لله الذي بنعمته تتمّ الصالحات، وبتوفيقه يُنال النجاح وتُذلل الصعوبات. أحمده سبحانه وأشكره على ما منّ به عليّ من عون وتيسير في إنجاز هذه الدراسة، وأسال الله العلي العظيم أن يجعل هذا العمل خالصًا لوجهه الكريم، وأن ينفع به طلبة العلم والمهتمين بهذا المجال.

ولا يسعُني إلا أن أتقدّم بخالص الشكر وعظيم الامتنان لمشرفي الفاضل الدكتور شريف بن عبد الرحمن السعودي على ما قدّمه لي من دعم وتوجيهات علمية قيّمة كان لها الأثر البالغ في إنجاز هذا العمل، وأتقدّم له بالشكر الجزيل أيضًا على ما أفاض به عليّ من معرفة علمية وتشجيع وتوجيه طوال مشواري الدراسي. وأوجّه شكري الجزيل كذلك لأساتذتي الكرام في قسم علم النفس بكلية الآداب والعلوم الإنسانية في جامعة الشرقية، الذين كانوا مشاعل علم ونور، فقد كان لتوجيهاتهم ودروسهم الأثر الكبير في مسيرتي الدراسية.

ولا يفوتني أن أتوجّه بالشكر والتقدير إلى المختصّين في مركز القياس والتقويم بوزارة التربية والتعليم على تعاونهم الكريم وإتاحتهم البيانات اللازمة لإكمال هذه الدراسة، فقد كان لتعاونهم أثر عظيم في إنجاز هذا الدراسة. وأتقدّم بجزيل الشكر وعاطر الثناء للأفاضل أعضاء لجنة المناقشة الكرام على ما تفضّلوا به من وقت وجهد في الاطلاع على هذا العمل، وعلى ملاحظاتهم القيمة التي أسهمت في تطوير هذا العمل وتحسينه.

وكذلك أعبّر عن بالغ امتناني لزملائي في العمل على دعمهم المستمرّ ومساندتهم الصادقة التي كانت عونًا لي في مراحل إعداد هذه الرسالة.

الباحث

ملخص الدراسة

معادلة درجات امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة في سلطنة عُمان باستخدام نظرية الاستجابة للفقرة

الباحث: علي بن عامر بن سلّوم الهنائي

المشرف: د. شريف بن عبد الرحمن السعودي (مشرفًا رئيسًا)، د. عامر بن سالم الحبسي (مشرفًا ثانيًا)

هدفت الدراسة إلى معادلة درجات صورتين من امتحان الرياضيات المتقدمة لدبلوم التعليم العام بسلطنة عُمان باستعمال نظرية الاستجابة للفقرة عن طريق توظيف تصميم المجموعات العشوائية (المتكافئة) ووضعهما على تدريج مشترك، حيث استعملت الدراسة صورتين لامتحان الرياضيات المتقدمة لدبلوم التعليم العام، وتمثّلت الصورة الأولى في امتحان الفصل الدراسي الأول للعام الدراسي 2023/ 2024م، وتمثّلت الصورة الثانية في امتحان الفصل الدراسي الأول للعام الدراسي 2024/ 2025م، واشتملت صورتا الامتحان على (27) فقرة و(29) فقرة على التوالي تنوّعت بين الفقرات ثنائية الاستجابة ومتعدّدة الاستجابة. وقد بلغ عدد الطلبة الذين تقدموا لامتحان مادة الرياضيات المتقدمة لدبلوم التعليم العام في العامين الدراسيين: 2023/ 2024م و2024/ 2025م (24619) طالبًا وطالبة و(28659) طالبًا وطالبة على التوالي. وُظِّفَ برنامج R حزمة (mirt) لتقدير معالم فقرات صورتي الامتحان، وتقدير قدرات الأفراد، فقد استعمل النموذج ثلاثي المعلمة (3PLM) لمعايرة الفقرات ثنائية الاستجابة ونموذج التقدير الجزئي المعمم (G–PCM) لمعايرة الفقرات متعددة الاستجابة، ثم أُنْجِزَت عملية المعادلة باستعمال ثلاث طرائق من طرائق نظرية الاستجابة للفقرة (طريقة معادلة الدرجات الخام، وطريقة معادلة الدرجات الحقيقية، وطريقة معادلة درجات القدرة) بتوظيف برنامج R. وقد أوضحت نتائج المعادلة باستعمال الدرجات الخام أن امتحان الصورة الأولى كان أسهل نسبيًّا؛ حيث إن الدرجة الخام على الصورة الأولى ارتبطت بدرجة خام أقل منها أو تساويها على الصورة الثانية، وأوضحت نتائج معدلة الدرجات

الحقيقية تقارب الإحصاءات الوصفية للدرجات الحقيقية على صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام، حيث توزعت الدرجات الحقيقية على الصورة الأولى بمتوسط حسابي 44.61 وانحراف معياري 26.07، وتوزعت الدرجات الحقيقية على الصورة الثانية بمتوسط حسابي 43.01 وانحراف معياري 27.30، وأما بالنسبة لمعادلة درجات القدرة، حيث رُبِطَ بين الدرجة الخام والدرجة الحقيقية ودرجة القدرة، فقد أظهرت النتائج بشكل عام اختلاف الدرجة الخام على الصورة الأولى المقابلة لدرجة قدرة معينة عن الدرجة الحقيقية على الصورة الثانية المقابلة للقدرة نفسها.

الكلمات المفتاحية: معادلة درجات الامتحانات، وامتحان الدبلوم العام، والرياضيات المتقدمة، ونظرية الاستجابة للفقرة، والنموذج ثلاثي المعالم، ونموذج التقدير الجزئي المعمم.

Study Abstract

Equating Scores of the General Education Diploma Advanced Mathematics Exam in Oman Using Item Response Theory

Researcher: Ali Amer Sallom Al-Hinai

Supervisors: Dr. Sharif Abdulrahman Al-Saudi (Primary Supervisor),

Dr. Amir Salim Al-Habsi (Secondary Supervisor)

This study aimed to equate the scores of two versions of the General Education Diploma Advanced Mathematics exam in Oman using Item Response Theory (IRT). The study employed an equivalent groups (randomized) design to place both versions on a common scale. Two versions of the Advanced Mathematics exam were used: the first version was the first-semester exam for the academic year 2023/2024, and the second version was the first-semester exam for the academic year 2024/2025. The two exam versions consisted of 27 and 29 items, respectively, including both dichotomous and polytomous response items.

The number of students who took the Advanced Mathematics exam in the academic years 2023/2024 and 2024/2025 was 24,619 and 28,659, respectively. The R software package (mirt) was used to estimate item parameters for both exam versions and individual ability levels. The three-parameter logistic model (3PLM) was applied to calibrate dichotomous items, while the generalized partial credit model (G-PCM) was used for polytomous items. Equating was then conducted using three IRT-based methods: raw score equating, true score equating, and ability score equating, all implemented in R.

The results of raw score equating indicated that the first exam version was relatively easier, as a raw score on the first version corresponded to an equal or lower raw score on the second version. Meanwhile, true score equating revealed close descriptive statistics for true scores across both exam versions. True scores for the first version had a mean of 44.61 and a standard deviation of 26.07, while those for the second version had a mean of 43.01 and a standard deviation of 27.30.

For ability score equating, which linked raw scores, true scores, and ability scores, the results generally showed that a raw score on the first version corresponding to a given ability level differed from the true score on the second version for the same ability level.

Keywords: Test score equating, General Diploma Exam, Advanced Mathematics, Item Response Theory (IRT), Three-Parameter Logistic Model (3PLM), Generalized Partial Credit Model (G-PCM).

قائمة المحتويات

قرار لجنة المناقشة.....أ

الإقرار .....ب

إهـــداء .....ج

شكر وتقدير .....هـ

ملخص الدراسة باللغة العربية .....و

ملخص الدراسة باللغة الإنجليزية .....ح

قائمة المحتويات.....ط

قائمة الجداول .....م

قائمة الأشكال .....س

قائمة الملاحق.....ف

الفصل الأول: مشكلة الدراسة وأهميتها.....1

مقدمة.....2

مشكلة الدراسة وأسئلتها 6

أهداف الدراسة.....10

أهمية الدراسة.....10

الأهمية النظرية.....10

الأهمية العملية 11

محددات الدراسة.....12

مصطلحات الدراسة.....12

امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة 13

مناهج سلاسل العلوم والرياضيات (مناهج كامبردج).....13

نظرية الاستجابة للفقرة 14

الفصل الثاني: الإطار النظري والدراسات السابقة 15
أولا: الإطار النظري 16
الامتحان النهائي في مادة الرياضيات المتقدمة لدبلوم التعليم العام 16
نظرية الاستجابة للفقرة 17
نماذج نظرية الاستجابة للفقرة..... 19
النماذج ثنائية الاستجابة 20
النموذج اللوجستي أحادي المعلمة 20
النموذج اللوجستي ثنائي المعلمة 21
النموذج اللوجستي ثلاثي المعلمة 22
نماذج تحتوي على أكثر من فئتين من الدرجات 23
نموذج الاستجابة المتدرجة 24
نموذج الاستجابة المتدرجة المُعدل..... 27
نموذج التقدير الجزئي..... 27
نموذج التقدير الجزئي المعمم 29
نموذج مقياس التقدير 32
نموذج الاستجابة الاسمية 34
معادلة درجات الامتحانات 35
1- المعادلة الأفقية..... 37
2- المعادلة الرأسية 37
شروط (متطلبات) معادلة درجات الاختبارات..... 38
1- التماثل..... 39
2- قياس المحتوى نفسه..... 39
3- المساواة (التعادل) 39

4-اللاتغير (عدم التباين في المجتمع) 40
5- الثبات 40
تصاميم جمع البيانات لإجراء المعادلة 40
1- تصميم المجموعات العشوائية أو المتكافئة 41
2 - تصميم المجموعة الواحدة 42
3- تصميم المجموعة الواحدة مع التوازن 43
4- تصميم الجذع المشترك للمجموعات غير المتكافئة 43
5- تصميم الجذع المشترك 44
طرائق معادلة درجات الاختبارات 44
1- طرائق معادلة الدرجات باستعمال النظرية الكلاسيكية..... 45
طريقة معادلة الوسط الحسابي 45
المعادلة بواسطة الرتب المئينية 46
2- طرائق المعادلة وفقًا لنظرية الاستجابة للفقرة 46
معادلة الدرجات باستعمال الدرجات الحقيقية 47
معادلة الدرجات باستعمال درجات القدرة 48
معادلة الدرجات باستعمال الدرجات الملاحظة..... 48
طرائق حساب معاملات تحويل المقاييس (معاملات الربط) 48
ثانيًا: الدراسات السابقة..... 51
مناقشة الدراسات السابقة 61
الفصل الثالث: منهجية الدراسة وإجراءاتها 63
منهج الدراسة 64
مجتمع الدراسة وعينتها 64
أداة الدراسة 65

تصميم الدراسة 70
إجراءات الدراسة 70
المعالجات الإحصائية 72
الفصل الرابع: مناقشة النتائج 76
مقدمة 77
النتائج المتعلقة بالسؤال الأول ومناقشتها 77
النتائج المتعلقة بالسؤال الثاني ومناقشتها 86
النتائج المتعلقة بالسؤال الثالث ومناقشتها 88
النتائج المتعلقة بالسؤال الرابع ومناقشتها 90
النتائج المتعلقة بالسؤال الخامس ومناقشتها 101
النتائج المتعلقة بالسؤال السادس ومناقشتها 103
ملخص النتائج 115
التوصيات والمقترحات 116
المراجع 117
الملاحق 127

قائمة الجداول

الجدول (1) تصميم المجموعات العشوائية (المتكافئة) 42

الجدول (2) تصميم المجموعة الواحدة 42

الجدول(3) تصميم المجموعة الواحدة المتوازنة 43

الجدول (4) تصميم الجذع المشترك 44

جدول(5) أعداد المتقدمين لامتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة بسلطنة عُمان للعامين الدراسيين: 2023/ 2024م و2024/ 2025م وفق المحافظة والنوع الاجتماعي لصورتي الامتحان..... 65

جدول(6) توزيع درجات الامتحان النهائي لمادة الرياضيات المتقدمة حسب هدفي التقويم ومستوى الصعوبة والوحدة الدراسية..... 67

جدول (7) توزيع فقرات امتحان الصورة الأولى 2024/2023م 68

جدول (8) توزيع فقرات امتحان الصورة الثانية 2024/ 2025م..... 69

جدول (9) تصميم المعادلة المستعملة في الدراسة..... 70

جدول (10) الإحصاءات الوصفية لصورتي امتحان الرياضيات المتقدمة وقيم التفلطح والالتواء ... 77

جدول (11) التحقق من محكات صلاحية بيانات صورتي الامتحان للتحليل العاملي 78

جدول (12) نتائج التحليل العاملي لبيانات صورتي الامتحان..... 80

جدول(13) تشبعات فقرات صورتي الامتحان على العامل السائد لكل صورة 81

جدول (14) قيم معاملات ألفا كرونباخ وماكدونالد أوميجا 86

جدول (15) ارتباطات البواقي المعيارية لفقرات صورتي الامتحان 87

جدول(16) مؤشرات جودة مطابقة الفقرات ثنائية التصحيح في صورتي الامتحان للنموذج الثلاثي. 88

جدول(17) معالم فقرات امتحان الصورة الأولى 2023/ 2024م..... 91

جدول(18) معالم فقرات امتحان الصورة الثانية 2023/ 2024م 92

جدول(19) نسبة الأفراد المطابقين للنموذج بناء على مؤشر المطابقة Infit لصورتي الامتحان ...101

جدول (20) الإحصاءات الوصفية لتوزيع قدرات الأفراد 102
جدول (21) توزيع الأفراد على فئات القدرات لمجموعتي الأفراد.....103
جدول (22) الدرجات الخامّ على الصورة الأولى وما يقابلها على الصورة الثانية.....104
جدول (23) الإحصاءات الوصفية للدرجات المتعادلة على صورتي الامتحان والفروق بينها 105
جدول(24) توزيع درجات القدرة وما يقابلها من درجات حقيقية على صورتي الامتحان 108
جدول (25) الدرجات الخام على الصورة الأولى وما يقابلها من درجات قدرة ودرجات حقيقية على الصورة الثانية.....112
جدول (26) معامل ارتباط بيرسون بين الدرجات المحولة لصورتي الامتحان 113

قائمة الأشكال

شكل(1) منحنيات خصائص الفقرة في النموذج أحادي المعلمة 21

شكل(2) منحنى خصائص الفقرة في النموذج ثنائي المعلمة 22

شكل (3) منحنيات خصائص الفقرة في النموذج ثلاثي المعلمة 23

شكل (4) منحنيات خصائص الفئات لفقرة تتكون من خمس فئات وفق نموذج GRM 26

شكل (5) منحنيات خصائص استجابة الفئات لفقرة ذات خمس فئات استجابة وفق نموذج GRM. 26

شكل (6) منحنيات استجابة الفئات لفقرة مكونة من ثلاث فئات وفق نموذج PCM..... 29

شكل (7) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a1 = 0.5$ 30

شكل (8) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a2 = 1$..... 31

شكل (9) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a3 = 2$..... 31

شكل (10) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a3 = 2$ 31

شكل(11) منحنى استجابة الفئة للفقرة الأولى وفق نموذج (RSM) حيث $\lambda i = -0.444$ 33

شكل(12) منحنى استجابة الفئة للفقرة الثانية وفق نموذج (RSM) حيث $\lambda i = 0.333$ 33

الشكل (13) منحنيات استجابة الفئة لفقرة من خمس فئات استجابة وفق (NRM)..... 34

شكل (14) Scree plot للتحليل العاملي لبيانات الصورة الأولى 83

شكل (15) Scree plot للتحليل العاملي لبيانات الصورة الثانية..... 83

شكل(16) Scree plot للتحليل العاملي المتوازي لبيانات الصورة الأولى..... 84

شكل(17) Scree plot للتحليل العاملي المتوازي لبيانات الصورة الثانية 85

شكل (18) إحصائيات Infit وOutfit لصورة الامتحان الأولى 89

شكل (19) إحصائيات Infit وOutfit لصورة الامتحان الثانية..... 90

شكل(20) منحنى خصائص الفقرة لفقرات امتحان الصورة الأولى 2023/ 2024م..... 96

شكل(21) منحنى خصائص الفقرة لفقرات امتحان الصورة الثانية 2024/ 2025م 96

شكل(22) منحنيات دالة معلومات فقرات الصورة الأولى 2023/ 2024م..... 97

شكل(23) منحنيات دالة معلومات فقرات الصورة الثانية 2024/ 2025م 97
شكل (24) منحنى الدرجة الكلية المتوقعة لصورة الامتحان الأولى 2023/ 2024م 98
شكل (25) منحنى الدرجة الكلية المتوقعة لصورة الامتحان الثانية 2024/ 2025م 99
شكل (26) دالة معلومات الاختبار لصورة الامتحان الأولى 2023/ 2024م 100
شكل (27) دالة معلومات الاختبار لصورة الامتحان الثانية 2023/ 2024م 100
الشكل (28) الدرجات الخام المعادلة للصورة الثانية مع ما يقابلها من الدرجات الخام للصورة المرجعية 107
الشكل (29) مخطط الصندوق والمؤشر للدرجات الخام لصورتي الامتحان 107
الشكل(30) درجات القدرة وما يقابلها من درجات حقيقية على صورتي الامتحان 111

قائمة الملاحق

ملحق (1) رسالة طلب تسهيل مهمة باحث من جامعة الشرقية 127

ملحق (2) رسالة طلب بيانات من وزارة التربية والتعليم.....128

الفصل الأول

مشكلة الدراسة وأهميتها

- المقدمة
- مشكلة الدراسة وأسئلتها
- أهداف الدراسة
- أهمية الدراسة:
  - الأهمية النظرية
  - الأهمية العلمية
- محددات الدراسة
- مصطلحات الدراسة

الفصل الأول

مشكلة الدراسة وأهميتها

مقدمة:

يُعَدّ التقويم عنصرًا أساسيًّا من عناصر العملية التعليمية؛ فبواسطته يُحْكَم على فاعلية العملية التعليمية من حيث قدرتها على تحقيق الأهداف التربوية المنشودة.

ورغم تعدّد أدوات القياس التربوي إلا أن الاختبارات تُعَدّ إحدى أهمّ أدوات القياس في العملية التعليمية، حيث تهدف إلى تقييم قدرات الطلبة ومستوياتهم الأكاديمية وتحديد مدى تحقيقهم للأهداف التعليمية. وتحتلّ الاختبارات مكانةً محوريّةً في العملية التعليمية، حيث تكمن أهميتها في توفير مؤشرات واضحة حول مستوى الأداء الأكاديمي، مما يساعد على تحسين إستراتيجيات التعليم والتعلم، وتُوَظَّف الاختبارات كوسيلة أساسية لاتخاذ القرارات الأكاديمية المتعلقة بالنجاح أو الإخفاق، والانتقال بين المراحل الدراسية المختلفة، وتحديد مجالات التطوير لدى الطلبة.

فالامتحان التحصيليّ ليس عملًا ارتجاليًا، وإنما هو مجموعة من الإجراءات التي تمتاز بالتسلسل والتنظيم؛ فهو الطريقة المنظمة لتحديد كمية ما تعلّمه الطالب في مدّة زمنية معينة، أو في نهاية وحدة تعليمية، أو مقرر دراسيّ، ويختتم بمنح علامة تقديرية تعبّر عن أداء الطالب (بوصلب، 2023).

ومع تنامي دور الاختبارات في رسم المسارات التعليمية والمهنية، بات من الضروري ضمان موضوعية هذه الاختبارات وعدالتها، خصوصًا في ظلّ تنوّع البيئات التعليمية والفروق الفردية بين الطلبة. ونظرًا لضرورة المحافظة على سرية الاختبارات، ظهرت أهمية وجود أكثر من صورة اختبارية واحدة تقيس المحتوى نفسه؛ ولذا ظهرت أهمية معادلة درجات الاختبارات، فقد حظيت عملية معادلة درجات الاختبارات باهتمام كبير بين أوساط المختصين في القياس والتقويم التربويّ، حيث اهتمت الكثير من المؤسسات التربوية بمعادلة درجات الاختبارات، ومنها الجمعية الأمريكية لعلم النفس (العنزيّ، 2014). وعلى الرغم من أن

مطوري الاختبارات يبذلون جهدًا كبيرًا من أجل إعداد نماذج متكافئة من الاختبارات، إلا أنه من الصعب تحقيق هذا الهدف في كثير من الأحيان؛ لذا فإن النماذج المختلفة للاختبار تختلف على الأقل في درجة الصعوبة، والغرض من عملية المعادلة هو تعديل الاختلافات في الصعوبة (Kolen & Brennan, 2014)، فعند استعمال نماذج مختلفة من الاختبار لمجموعات مختلفة من الطلبة، فمن المتوقّع أن تتباين الدرجات بين النماذج المختلفة؛ فالذين أجابوا عن النموذج الأسهل سيحصلون على درجات ملاحظة مرتفعة، والذين أجابوا عن النموذج الأصعب سيحصلون على درجات ملاحظة منخفضة؛ ولذا فإن استعمال هذه الدرجات يؤدّي إلى قرارات خاطئة؛ لأن دقة القرار تعتمد على دقة المعلومات المتوافرة، وكذلك فإن الحصول على معلومات صادقة يُعَدّ المحور الأهمّ في عملية التقويم (عودة، 2010).

ويقصد بمعادلة نماذج الاختبار Test Form Equating عملية تحويل نظام وحدات القياس الخاصّ بأحد النموذجين إلى نظام وحدات النموذج الآخر، بحيث تصبح القياسات المستمدة من كلا النموذجين متكافئة بعد إجراء هذا التحويل (Kolen & Brennan,2014)، حيث تتطلّب هذه العملية تطبيق تصميم خاصّ يُسَمّى تصميم المعادلة Equating design، ويهدف تصميم المعادلة إلى ضبط مستوى القدرة والسمة التي يقيسها الاختبار بين العينات التي يطبق عليها الاختبار، بحيث يتسنّى أن نرجع الاختلاف في المتوسط الحسابي أو التوزيع التكراري لنماذج الاختبار إلى الاختلاف في مستوى صعوبتها، والذي يتطلّب القيام بعملية إحصائية على درجات أحد نماذج الاختبار للحصول على الدرجات المكافئة للنموذج الآخر (المحرزيّ، 2015).

وقد أشارت الزعابية وآخرون (2020) إلى أن الاهتمام من المختصّين في القياس والتقويم بموضوع معادلة الاختبارات بدأ متأخرًا، فخلال العشرين سنة المنصرمة بدأ الاعتراف بأهمية معادلة الاختبارات؛ فقد بحثوا فيه بشكل مستفيض اعترافًا بأهميته في بناء الاختبارات، ويمكن أن يعزى هذا الاهتمام المتزايد بموضوع الاختبارات إلى تصورات مهمّة، حيث لخّصها الدوسريّ (2001) في ثلاث تصوّرات: فالتصوّر

الأول يهتمّ بزيادة عدد برامج الاختبارات وتنوّعها؛ مما يؤدّي إلى ضرورة معادلة الاختبارات التي تقيس السمة نفسها، والتصوّر الثاني يتعلّق بمطوري الاختبارات الذين أدركوا أهمية معادلة درجات الاختبارات لمعالجة الكثير من القضايا الفنية التي يثيرها النقاد عند تحليل الدرجات، بينما يتعلّق التصوّر الثالث بقضية قدرة الاختبارات على العدل في قياس السمة لمجموعة مختلفة من الأفراد المختلفين في الخصائص.

وتُعَدّ عملية معادلة درجات الاختبارات من المجالات التطبيقية لنظرية الاستجابة للفقرة، حيث تسمح لنا بتحويل درجات النماذج المختلفة للاختبار إلى مقياس مشترك موحّد يمكننا به إجراء المقارنات بين الدرجات، حيث إنه من دون إجراء عملية المعادلة تكون عملية المقارنة غير عادلة بسبب صعوبة فقرات أحد النماذج أو سهولتها؛ لذلك فلا بدّ من تحويل الدرجات الخامّ التي نحصل عليها من نموذجين مختلفين لاختبار معين، بحيث تصبح الدرجات المحوّلة لكل منهما على مقياس مشترك موحّد (العنزيّ، 2021).

ولإجراء معادلة الدرجات باستعمال نظرية الاستجابة للفقرة فهناك ثلاث طرائق هي المعادلة باستعمال الدرجات الحقيقية True-Score Equating، أو المعادلة باستعمال درجات القدرة Ability Score Equating، أو المعادلة باستعمال الدرجات الملاحظة Observed Score Equating (Kolen & Brennan,2014)، ويوجد نوعان من أنواع المعادلة هما المعادلة الأفقية (Horizontal equating)، والمعادلة الرأسية (Vertical equating)، كذلك فإن هناك الكثير من التصميمات الإحصائية التي تستعمل بغرض إجراء معادلة صور الاختبارات، فإما أن نستعمل تصميم المجموعة الواحدة، أو تصميم المجموعات المتكافئة أو العشوائية، أو تصميم المجموعات العشوائية المتوازنة، أو تصميم المجموعات غير المتكافئة ذات الجذع المشترك، أو نستعمل تصميم المجموعات المتكافئة ذات الجذع المشترك، وتشير دقة المعادلة إلى مدى دقة عملية المعادلة وصحّتها، والتي تقاس بإيجاد الخطأ المعياريّ لعملية المعادلة (العنزيّ، 2014).

وتحتل مرحلة دبلوم التعليم العامّ مكانة محورية بين مراحل التعليم المدرسيّ؛ كونها تمثّل المرحلة النهائية من مراحل التعليم المدرسيّ، وتُعَدّ البوابة الرئيسة للانتقال إلى التعليم العالي، حيث تُمثّل هذه المرحلة الأساس الذي يعتمد عليه الطالب في اتخاذ قرارات مصيرية تتعلّق بمستقبله، سواء كان ذلك في اختيار التخصّص الجامعيّ أو دخول سوق العمل، وإلى جانب ذلك تُعَدّ مرحلة دبلوم التعليم العامّ مؤشرًا على مدى نجاح النظام التعليميّ في تحقيق أهدافه، حيث تعكس نتائجها مستوى جودة التعليم المدرسيّ ومدى توافقه مع احتياجات العصر.

وتقوم وزارة التربية والتعليم في سلطنة عُمان بإعداد امتحانات دبلوم التعليم العامّ، حيث يقوم فريق مختصّ في المادّة العلمية ببناء امتحانات تتضمّن أسئلة موضوعية وأسئلة مقالية (الحبسي، 2018)، ويقوم فريق مختصّ في وزارة التربية والتعليم بإعداد الامتحانات النهائية وفق جداول المواصفات المبنية والموضّحة في وثيقة التقويم الخاصّة بكلّ مادة دراسية.

وتُعَدّ مادة الرياضيات إحدى الركائز الأساسية في العلوم الحديثة، حيث تمثّل أداة حيوية لفهم العالم من حولنا وحلّ المشكلات المعقّدة في المجالات المختلفة، حيث تتجاوز هذه المادة العمليات الحسابية البسيطة، لتغطّي مفاهيم رياضية متعمّقة تشمل الجبر المتقدّم، والتفاضل والتكامل، والهندسة التحليلية، والإحصاء، وغيرها من المفاهيم الرياضية؛ مما يجعلها ضرورية لتطبيقات متنوّعة في الهندسة، والفيزياء، والاقتصاد، وعلوم الحاسوب، والبحث العلميّ. ومن منطلق أهمية الرياضيات في العلوم الحديثة سعت وزارة التربية والتعليم إلى تطبيق مناهج عالمية بالتعاون مع مؤسسة كامبريدج، حيث تقوم هذه المناهج على دمج عملية التعلم مع مهارات المستقبل (وزارة التربية والتعليم،2022أ). ويُعَدّ العام الدراسي 2024/ 2025م هو العام الدراسي الثاني لتطبيق مناهج سلاسل كامبريدج في الصف الثاني عشر،

ويشتمل منهج الرياضيات المتقدمة للصف الثاني عشر للفصل الدراسي الأول على موضوعات: القياس الدائري– حساب المثلثات– مقدمة في النهايات والاتصال والتفاضل (وزارة التربية والتعليم،2024أ).

وجاءت هذه الدراسة لمعادلة درجات مادة الرياضيات المتقدمة لدبلوم التعليم العام في سلطنة عُمان وللتعريف بماهية معادلة درجات الامتحانات وطبيعتها وكيفية استفادة القائمين على الامتحانات بوزارة التربية والتعليم من معادلة درجات الامتحانات.

مشكلة الدراسة وأسئلتها

تحرص مؤسسات إعداد الامتحانات ومراكزها، ومنها مركز القياس والتقويم بوزارة التربية والتعليم في سلطنة عُمان على اتباع خطوات علمية عند إعداد الامتحانات وتصحيحها– ومنها امتحانات دبلوم التعليم العام– التي تجد اهتمامًا كبيرًا من الطلبة وحرصًا منهم على الحصول على أعلى درجات ممكنة؛ لما يترتّب عليها من آثار على مستقبلهم العلميّ والوظيفيّ، حيث يشير الحبسي (2018) إلى أن إعداد الامتحانات يتطلّب بذل الكثير من الجهد والوقت، ويواجه المختصّون في إعداد الاختبارات وتطويرها مشكلات كثيرة، منها ما يتعلّق بعدالة الامتحانات للفئات المستهدفة وخاصةً عند اختلاف مواعيد تطبيق الامتحانات؛ لذا فلا بدّ أن تتصف هذه الامتحانات بالعدالة والصدق والموضوعية في قياس تحصيل الطلبة الذي يعتمد عليه قَبولهم في مؤسسات التعليم العالي. ويشير Dorans et al. (2010) إلى أن معادلة درجات الاختبارات تعمل على إزالة التأثيرات الناتجة عن الاختلاف غير المقصود في صعوبة أحد النماذج، حيث تُعَدّ معادلة درجات الامتحان ضرورية لتحقيق العدالة بين الطلبة، وهي توفّر للقائمين على الاختبار بيانات تمكّنهم من المقارنة بين الصور المختلفة وكذلك بين المجموعات المختلفة للطلبة، فقد تختلف صعوبة الاختبار من نموذج إلى آخر، وكذلك فإنه عادةً لا تستعمل الفقرات نفسها في النماذج المختلفة؛ وذلك للحفاظ على سرية الاختبار؛ ولذا فإن قَبول الطالب في مؤسسات التعليم العالي لا يتأثر

بقدرة الطالب فقط، ولكن سيتأثر أيضًا بصعوبة نماذج الاختبار (Wallin & Wiberg, 2025)، وبعبارة أخرى يتمثّل دور عملية المعادلة في ضمان الحصول على قياس دقيق لمهارات المتقدمين للاختبار بغضّ النظر عن النموذج الذي تعرّضوا له (von Davier, 2011).

فعندما يُطَبَّق الاختبار بشكل متكرّر، تُعَدّ نماذج متعددة من الاختبار؛ لمنع كشف محتوى الاختبار، ففي أغلب الأحيان لا تكون هذه النماذج متكافئة في محتواها ومستوى صعوبتها؛ ولذا سيكون للمفحوصين الذين يأخذون الاختبار السهل أفضلية على الطلبة الذين سيأخذون الاختبار الصعب، لكن إذا عُودِلَت درجات نماذج الاختبارات بطريقة علمية صحيحة، فإن نموذج الاختبار الذي تعرّض له المفحوص لا يؤثّر في درجته؛ ولذا سيحصل المفحوصون من القدرة نفسها على الدرجات نفسها (المحروق، 2016).

وتشير De Ayala (2009) إلى أنه من غير الممكن عمليًّا أن نبنيَ صورًا متكافئة بصورة تامة للاختبار الواحد؛ وذلك بمراعاة الخصائص الإحصائية للاختبار، بالإضافة إلى أن ظروف التطبيق قد تؤثر في أداء الأفراد على صور الاختبار؛ فتتأثر تبعًا لذلك الخصائص الإحصائية للصور المختلفة للاختبار، وفي هذه الحالة فإن معادلة درجات الاختبار تمكّننا بدرجة مقبولة من التخلّص من الفروق بين هذه الصور؛ ولذا نتمكّن من المقارنة بين أداء الأفراد بغضّ النظر عن الصورة الاختبارية المقدّمة لهم.

فهناك حاجة ماسة لمعادلة درجات الاختبارات من أجل تحقيق نزاهة الدرجات والبيانات التي يُحْصَل عليها وصدقها، وأيضًا فإنها تحول دون الاضطرار إلى إعادة الاختبار أكثر من مرة، حيث تُعَدّ عملية معادلة درجات الاختبار وسيلة مهمة لضمان العدل في توزيع الأفراد والقضاء على تأثير العوامل المختلفة في الدرجات (Felan, 2002).

وتُعَدّ امتحانات الدبلوم العام في سلطنة عُمان إحدى أدوات التقويم الختامية، والتي تُنَفَّذ في نهاية كلّ فصل دراسي؛ لتحديد مستويات الطلبة، ومعرفة مدى تحقّق الأهداف التعليمية للمحتوى الدراسيّ، حيث يقوم فريق مختصّ من وزارة التربية والتعليم بإعدادها في كلّ عام دراسيّ، وفي كلّ عام دراسيّ

يختلف نموذج الاختبار عن امتحانات الأعوام السابقة، ولكنه لا يوجد دليل على أن هذه النماذج من الامتحانات متكافئة في صعوبتها؛ ولذا لا يمكن إجراء مقارنة علمية بين مستويات الطلبة في الأعوام الدراسية المختلفة (الحبسي، 2018)؛ حيث إن تشابه المحتوى المقاس باختبارات قد بُنِيَت من جدول المواصفات نفسه لا يضمن تكافؤ الصعوبة بين هذه النماذج المختلفة.

ونظرًا لوجود صورة مختلفة في كلّ عام دراسيّ لامتحانات دبلوم التعليم العامّ في مادة الرياضيات المتقدمة، جاء اختيار الباحث لامتحان مادة الرياضيات المتقدمة؛ لخبرته العملية في تدريس مادة الرياضيات، حيث إنه عند تحويل الدرجات الخامّ للطلبة إلى درجات متعادلة باستعمال نموذجين مختلفين لامتحان مادة الرياضيات المتقدمة يقيسان المحتوى نفسه، ويبنيان بجدول المواصفات نفسه، نتمكّن من المقارنة بين درجات الطلبة في الصورتين، وكذلك نستطيع تقليل الأخطاء عند تفسير نتائج الطلبة، وكذلك فإن دراسة الحبسي (2018) أوصت بإجراء دراسات لمعادلة درجات امتحانات مواد دراسية أخرى.

وعلى الرغم من الجهود التي يبذلها القائمون على إعداد امتحانات الدبلوم العام بوزارة التربية والتعليم، وحرصهم على بناء اختبارات متكافئة في صعوبتها، إلا أن هناك عوامل تحول دون تحقّق ذلك بصورة عملية، منها طبيعة امتحانات الدبلوم العام، وحساسية المرحلة التعليمية وأهميتها، فكلّ ذلك يتطلّب منهم تجنّب تكرار الفقرات الامتحانية، وكذلك تجنّب تكرار أفكار الفقرات الامتحانية؛ فيؤدّي ذلك إلى وجود فرق في صعوبة الصور الامتحانية المختلفة.

وأيضًا فإن ملاحظات الطلبة وشكواهم حول وجود فرق في صعوبة الصور الاختبارية في الأعوام المختلفة تشكّل هاجسًا وقلقًا للمختصّين في مركز القياس والتقويم، وهي تشكّل كذلك قلقًا لأولياء أمور الطلبة خوفًا على مستقبل أبنائهم؛ فلذا تُعَدّ معادلة الامتحانات هي الطريقة العلمية التي يمكن بها من التحقّق من صحة ملاحظات الطلبة حول صعوبة بعض الصور الاختبارية، والتحقّق من موضوعية هذه الصور الاختبارية وعدالتها.

وتُعَدّ الدراسة الحالية هي الدراسة الأولى لمعادلة امتحانات مناهج كامبريدج بشكل عامّ ومادة الرياضيات المتقدمة بشكل خاصّ، حيث تهدف إلى معادلة درجات امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة في سلطنة عُمان، وتعديل الفرق في الصعوبة بين الصورتين المختلفتين باستعمال نظرية الاستجابة للفقرة، ووضع صورتي امتحان مادة الرياضيات المتقدمة للعامين الدراسيين: (2023/ 2024م) و(2024/ 2025م) على تدريج مشترك. وستحاول الدراسة الحالية الإجابة عن الأسئلة الآتية:

- ما دلالات صدق البناء والثبات لصورتي الامتحان النهائيّ للفصل الدراسيّ الأول في مادة الرياضيات المتقدمة لدبلوم التعليم العامّ للعامين الدراسيين: 2023/ 2024م، و2024/ 2025م؟
- ما مدى مطابقة درجات صورتي الامتحان النهائيّ للفصل الدراسي الأول لمادة الرياضيات المتقدمة لدبلوم التعليم العامّ لافتراضات نظرية الاستجابة للفقرة؟
- ما مدى مطابقة فقرات صورتي الامتحان النهائيّ للفصل الدراسي الأول في مادة الرياضيات المتقدمة لنموذج نظرية الاستجابة للفقرة (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و (نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية؟
- ما تقدير معالم فقرات صورتي الامتحان النهائيّ للفصل الدراسي الأول في مادة الرياضيات المتقدمة باستعمال (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية؟
- ما مدى مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة وما تقدير قدرات الأفراد؟
- ما قيم الدرجات المتعادلة لصورتي الامتحان النهائي لمادة الرياضيات المتقدمة لدبلوم التعليم العامّ؟

أهداف الدراسة

تهدف الدراسة الحالية إلى معادلة درجات صورتين من امتحان الرياضيات المتقدمة لدبلوم التعليم العامّ بسلطنة عُمان باستعمال نظرية الاستجابة للفقرة عن طريق توظيف تصميم المجموعات العشوائية (المتكافئة)، وتحقيق الأهداف الآتية:

- تحديد خصـــائص توزيع الدرجات الخامّ والتحقّق من دلالات صـــدق البناء والثبات لصـــورتي الامتحـان النهـائيّ لمـادة الريـاضـــيـات المتقدمـة لـدبلوم التعليم العـامّ، والتحقق من مطـابقتهـا لافتراضات نظرية الاستجابة للفقرة.
- التحقّق من مطابقة فقرات صـــورتي الامتحان النهائي للفصـــل الدراســـي الأول في مادة الرياضـــيات المتقدمة لنموذج نظرية الاستجابة للفقرة (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية)، وتقدير معالمها باستخدام النموذجين.
- التحقّق من مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة وتقدير قدرات الأفراد.
- إيجاد قيم الدرجات المتعادلة لصورتي الامتحان النهائيّ لمادة الرياضيات المتقدمة لدبلوم التعليم العامّ.

أهمية الدراسة

تتمثّل أهمية هذه الدراسة في استعمالها لنظرية الاستجابة للفقرة لمعادلة درجات امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستخدام الدرجات الملاحظة (الخامّ)، والدرجات الحقيقية (المتوقعة)، ودرجات القدرة، ووضعها في تدريج مشترك، ويمكن تلخيص أهمية الدراسة في الجوانب الآتية:

الأهمية النظرية:

تتمثّل الأهمية النظرية للدراسة في النقاط الآتية:

- توفر معلومات للباحثين المهتمّين بعملية معادلة درجات الاختبارات، وطرائقها، وتصـــاميم جمع البيانات؛ مما يساعدهم في عملية إجراء المعادلة، ويؤدّي إلى تحسين جودة الامتحانات المقدمة للطلبة.

- تُسْــهم في توجيه أنظار المختصّــين بمركز القياس والتقويم إلى أهمية معادلة درجات الاختبارات في عملية المقارنة بين مستويات الأداء للطلبة في الأعوام المختلفة.
- تســاعد المختصّــين والباحثين بموضــوع معادلة درجات الاختبارات على اختيار طريقة المعادلة المناسبة لطبيعة الاختبار؛ وذلك للحصول على نتائج دقيقة وواضحة.
- تُسْهم في توصيف خطوات استعمال نظرية الاستجابة للفقرة وإجراءاتها في معادلة درجات امتحانات الدبلوم العام؛ حيث تُعَدّ هذه الدراســـة الثانية من نوعها في مجال معادلة امتحانات الدبلوم العام في سلطنة عُمان، والأولى في مجال امتحانات السلاسل العالمية (مناهج كامبريدج).

الأهمية العملية:

- توفير معلومات وبيانات تســاعد المختصّــين في مركز القياس والتقويم حول تكافؤ الصــور المختلفة لامتحانات الدبلوم العامّ لمادة الرياضيات المتقدمة؛ ولذا فهي تُسْهم في تطوير عملية بناء الاختبارات.
- تساعد المختصين والمهتمين بمعادلة درجات الامتحانات عن طريق إعطائها صورة واضحة لهم في كيفية معادلة درجات الاختبارات.
- تُسْهم في تقديم أساس عمليّ يمكن الاستعانة به في بناء اختبارات ذات صور مختلفة متكافئة أو متعادلة، وتمتاز بخصائص سيكومترية جيدة.
- اســـتفادة المختصـــين بمركز القياس والتقويم من نتائج الدراســـة في إجراء مقارنة عملية بين مستويات الطلاب.
- استفادة المختصين بمركز القياس والتقويم من نتائج الدراسة لإعداد بنوك أسئلة وبنائها، بحيث تحتوي على فقرات محددة المعالم.

محددات الدراسة

تقتصر الدراسة على:

- درجات الطلبة في امتحان الدبلوم العامّ لمادة الرياضـــيات المتقدمة في ســـلطنة عُمان للعامين الدراسيين: (2024/2023م) والعام (2025/2024م).
- اســتعمال النموذج اللوجســتيّ الثلاثيّ لنظرية الاســتجابة للفقرة **(3PLM)** لمعايرة معالم الفقرات الموضـــوعية وتقديرها، واســـتعمال نموذج التقدير الجزئيّ المعمّم **Generalized Partial Credit Model (G−PCM)** لمعايرة معالم الفقرات المقالية وتقديرها لصورتي الامتحان.
- استعمال تصميم المجموعات العشوائية المتكافئة لجمع البيانات.
- استعمال حزمة **(mirt)** من برمجية R لتقدير معالم الفقرات وقدرات الأفراد.
- توظيف برمجية R لإجراء المعادلة لدرجات صورتي امتحان مادة الرياضيات المتقدمة.
- تعتمد دقة النتائج على دقة تقدير البرمجيات المستعملة في هذه الدراسة.

مصطلحات الدراسة

- **معادلة درجات الامتحانات (Equating Test Scores)**: هي عملية إحصـائية تُوضـع عن طريقها درجات الصور المختلفة للاختبارات التي تقيس السمة (المحتوى) نفسها على تدريج مشـترك؛ حتى تصـبح الدرجات متكافئة، ويمكن اسـتعمالها بصـورة متبادلة (الحربيّ وعليان، 2009)، ويُعرّفها Kolen and Brennan (2014) بأنها عملية إحصـــائية تســـتعمل لموائمة الدرجات في الصـــور المختلفة للاختبار، حيث يمكن اســـتعمال الدرجات بين الصـــور بالتبادل، حيث تعمل على تعديل الاختلافات في الصـــعوبة بين الصـــور؛ لتكون متشـــابهة في الصعوبة والمحتوى.

وفي هذه الدراسة يقصد بمعادلة درجات امتحان مادة الرياضيات المتقدمة بأنها تحويل نظام درجات امتحان الرياضيات المتقدمة لعام (2024/2023م) إلى نظام درجات امتحان الرياضيات المتقدمة لعام (2025/2024م)، وذلك بحساب الدرجات الخامّ لمجموعة الإجابات الصحيحة لفقرات إحدى الصورتين وتحويل درجات إحدى الصور إلى الصورة المرجعية التي اعتمدها الباحث، وهي امتحان العام الدراسي (2024/2023م).

- **امتحان دبلوم التعليم العامّ لمادة الرياضيات المتقدمة:** هو أداة تقويم ختامية يُعَدّ على مستوى وزارة التربية والتعليم، ويُطَبَّق في نهاية الفصل الدراسي، ويراعى عند إعداده الأوزان النسبية للمخرجات التعليمية للمنهج والمستويات التعليمية لها، حيث توجد ورقة امتحانية لكلّ فصل دراسي، وحُدِّدَت الدرجة الكلية للامتحان النهائيّ في مادة الرياضيات المتقدمة لدبلوم التعليم العام بـ 70 درجة (وزارة التربية والتعليم، 2024ب).
- **مناهج سلاسل العلوم والرياضيات (مناهج كامبردج):** هي مناهج عالمية اختيرت من المختصّين بوزارة التربية والتعليم بناءً على رؤية واضحة وبمعايير دقيقة، إذ تشير الإحصائيات إلى أن هناك أكثر من (160) دولة حول العالم تطبّق هذه المناهج بواقع (10,000) مدرسة، وأن 93 % من خريجي هذه المدارس قد قُبِلوا في الجامعات، ومنهم 70 % قُبِلوا في أفضل 500 جامعة في العالم، حيث تعترف بمؤهلات كامبردج الدولية أكثر من (2000) مؤسسة تعليمية بما في ذلك أفضل (20) جامعة في العالم، وأن جميع مراكز قَبول الطلبة في الجامعات المختلفة تقرّ بأن مناهج كامبردج تُعِدّ الطالب للجامعات بشكل متميز؛ إذ تعتمد على الاستقصاء العلمي وحلّ المشكلات بما يمكّن الطلبة من تنمية مهارات المستقبل ومهارات التفكير العلميّ لديهم. وتتبع هذه المناهج المبدأ الحلزونيّ في عرض الموضوعات؛ مما يجعل الطالب يتدرّج في الوصول إلى عمق المعلومات العلمية والمفاهيم بما يتناسب مع المرحلة العمرية (وزارة التربية والتعليم، 2022ب).

- **نظرية الاســـتجابة للفقرة:** هي عبارة عن مجموعة من الدوال الرياضـــية الاحتمالية التي تحدّد العلاقة بين اسـتجابات الأفراد على إحدى فقرات الاختبار وقدرات الأفراد، والتي يفترض أنها هي الســبب لاســتجابات الأفراد، وذلك باســتعمال (بارمتر) أو معالم الفقرات كالصــعوبة والتمييز والتخمين، حيث إنه كلما زادت قدرة الفرد، زاد احتمال إجابته عن الفقرة إجابة صــحيحة (جاد الله وآخرون، 2021). وتستند هذه النظرية إلى مجموعة من الافتراضات، وسوف نتطرق إليها بشكل مفصل في الفصل الثاني من هذه الدراسة.

الفصل الثاني

الإطار النظري والدراسات السابقة

أولًا. الإطار النظري:

- الامتحان النهائي في مادة الرياضيات المتقدمة لدبلوم التعليم العام
- نظرية الاستجابة للفقرة
  - افتراضات نظرية الاستجابة للفقرة
  - نماذج نظرية الاستجابة للفقرة
- معادلة درجات الاختبارات
  - مفهوم معادلة درجات الاختبارات
  - أنواع المعادلة
  - متطلبات المعادلة
  - تصاميم جمع البيانات
  - طرائق المعادلة حسب النظرية الكلاسيكية للقياس
  - طرائق المعادلة حسب نظرية الاستجابة للفقرة

ثانيًا: الدراسات السابقة

الفصل الثاني

الإطار النظري والدراسات السابقة

أولًا. الإطار النظري:

يتناول الإطار النظري لهذه الدراسة معلومات عن الامتحان النهائيّ في مادة الرياضات المتقدمة لدبلوم التعليم العامّ، ويتناول كذلك خلفية معرفية لنظرية الاستجابة للفقرة عن طريق توضيح افتراضاتها ونماذجها اللوجستية، وسيتناول أيضًا موضوع معادلة درجات الامتحانات، حيث سيُوَضَّح مفهومها وأنواعها وتصميماتها، وكذلك متطلبات إجراء معادلة درجات الامتحانات، وسيكون هناك تطرّق إلى طرائق معادلة درجات الامتحانات باستعمال نظريتي القياس النظرية الكلاسيكية ونظرية الاستجابة للفقرة.

الامتحان النهائيّ في مادة الرياضيات المتقدمة لدبلوم التعليم العامّ

يقوم مجموعة من المختصّين في المادة العلمية بإعداد امتحان الرياضيات المتقدمة لدبلوم التعليم العامّ وفق جدول مواصفات حدّدته وثيقة تقويم تعلّم الطلبة لمادة الرياضيات (2024ب)، حيث تتنوّع الفقرات التي يمكن استعمالها في الاختبار النهائي كما حدّدته وزارة التربية والتعليم في وثيقة تقويم الطلبة لمادة الرياضيات (2024ب)، ومنها:

1- فقرات الاختيار من متعدّد: يحتوي هذا النوع من الفقرات على مجموعة من الإجابات يختار الطالب إحدى هذه الإجابات، ويمكن للطالب الإجابة عن هذه الفقرات في مدّة زمنية قصيرة باسترجاع المعلومات أو تحليل البيانات المعطاة.

2- فقرات الرسم وتعيين النقاط: يشتمل هذا النوع من الفقرات على الرسم أو تعيين النقاط على الرسم التخطيطيّ أو المحاور المجودة سابقًا أو الرسم في مساحة خالية.

3- الفقرات ذات الإجابة القصيرة: يتطلّب هذا النوع قيام الطالب باستعمال عبارات جبرية بسيطة، أو استرجاع المعلومات أو البحث من المعلومات الواردة في الجداول أو الرسوم البيانية المرفقة.

4- الفقرات ذات الإجابة الطويلة: تحتاج الإجابة عن هذا النوع وقتًا أطول، وتتطلّب مجموعة من الخطوات والمزيد من التفكير للتعامل مع هذه الفقرات والوصول إلى النتائج.

5- الفقرات ذات الإجابة المصاغة: يتطلّب الإجابة عن هذا النوع من الفقرات قيام الطالب بتقديم إجابة مصاغة بالكلمات، وذلك بإظهار قدرته على وصف المسائل والمفاهيم الرياضية وشرحها.

نظرية الاستجابة للفقرة Item Response Theory (IRT)

تُعَدّ نظرية الاستجابة للفقرة أحد التوجهات السيكومترية المعاصرة التي حاولت التغلب على الكثير من جوانب القصور في النظرية الكلاسيكية للقياس مثل اختلاف خصائص الفقرات باختلاف عينات المفحوصين، حيث إن خصائص فقرات الاختبار تعتمد على مستوى قدرات أفراد العينة وتوزيعها، وكذلك فإن الاختبارات في النظرية الكلاسيكية لا تزودنا بمعلومات دقيقة عن قدرات المفحوصين من ذوي القدرة العالية والمفحوصين من ذوي القدرات المنخفضة؛ فعند بناء الاختبارات على أساس النظرية الكلاسيكية يؤخَذ في الحسبان أن تتلاءم هذه الفقرات مع الأفراد ذوي القدرات المتوسطة.

ولتلافي هذه العيوب، ومن أجل تحقيق قياس أكثر موضوعية؛ حاول علماء القياس الاستفادة من التقدّم التكنولوجيّ في التوصّل إلى طرائق سيكومترية جديدة، حيث انبثقت عن هذه الطرائق الجديدة نماذج تُسَمّى نماذج السمات الكامنة أو نظرية الاستجابة للفقرة Item Response Theory، والتي تقوم على دالة احتمالية تربط بين معلمتين Tow Parameters هما قدرة المفحوص والأخرى تتعلّق بالفقرة، حيث تُقَدّر احتمالية إجابة المفحوص عن الفقرة إجابة صحيحة بالوصول إلى تقدير دقيق لهاتين المعلمتين (المحروق، 2016).

ومن مزايا نظرية الاستجابة للفقرة: تقدير قدرة الفرد مستقلًا عن عينة الفقرات التي تستعمل في تقدير القدرة Item Free، وكذلك تقدير معالم الفقرات مستقلًا عن قدرات عينة الأفراد Sample Free (Hambelton & Swaminathan, 1985).

ويذكر كلّ من (الزبون والصرايرة، (2017)؛ والشريفين والكيلاني، (2003)؛ وجاد الله وآخرون (2021)) أن نظرية الاستجابة للفقرة تفترض عددًا من الافتراضات في البيانات التي ستطبق عليها، ومن هذه الافتراضات:

1. **أحادية البُعد Unidimensionality:** وتعني أن فقرات الاختبار متجانسة، وتقيس السمة نفسها، أي أن جميع الفقرات يتطلّب حلّها النوع نفسه من الإجراءات والعمليات السلوكية، ولكنها تختلف فيما بينها في تدرّج صعوبتها.
2. **الاستقلال الموضعيّ Local Independence:** ويقصد به أن تكون أن تكون استجابات الفرد للفقرات المختلفة في الاختبار مستقلّة إحصائيًّا عند مستوى قدرة معين، بمعنى أن استجابة الفرد لإحدى الفقرات لا تؤثّر سلبًا أو إيجابًا في استجابته للفقرات الأخرى من الاختبار.
3. **منحنى خصائص الفقرة Item Characteristic Curve (ICC):** وهو عبارة عن دالة رياضية تربط بين احتمال إجابة الفرد عن الفقرة إجابة صحيحة والقدرة التي تقيسها هذه الفقرة عن طريق اقتران تراكمي صاعد، حيث يوفر هذا المنحنى احتمالات إجابة المفحوصين عن الفقرة في مستويات القدرة المختلفة إجابة صحيحة، حيث إن احتمال إجابة الفرد عن الفقرة إجابة صحية يزداد بزيادة مستوى القدرة، وتتفق جميع نماذج نظرية الاستجابة للفقرة على وجود هذا المنحنى، لكنه يختلف من نموذج إلى آخر باختلاف عدد معالم (بارامتر) الفقرات، ففي النموذج الأحادي يعتمد على بارامتر واحد (معلمة واحدة) هو معلمة الصعوبة، وفي النموذج الثنائي يعتمد المنحنى على بارامترين

(معلمتين) هما معلمتا الصعوبة والتمييز، وفي النموذج الثلاثي يعتمد على ثلاثة بارامترات (معالم) هي معلمة الصعوبة ومعلمة التمييز ومعلمة التخمين.

4. **التحرّر من السرعة:** حيث تفترض نماذج نظرية الاستجابة للفقرة أن عامل السرعة لا يلعب دورًا في الإجابة عن فقرات الاختبار، أي أن عدم إجابة الفرد على الفقرة يعود إلى انخفاض قدرته، وليس إلى تأثير عامل السرعة في إجاباته، وأن الاختبار الذي يعتمد على السرعة يخلّ بافتراض أحادية البُعد.

نماذج نظرية الاستجابة للفقرة:

تعتمد نظرية الاستجابة للفقرة على مجموعة من النماذج الرياضية التي تهدف إلى تحديد العلاقة بين أداء الفرد في الاختبار والسمة التي تكمن خلف هذا الأداء وتفسّره، وتظهر هذه النماذج في صورة دوال رياضية تربط بين احتمال إجابة الفرد عن الفقرة إجابة صحيحة وبين مستوى قدرته (Hambelton & Swaminathan, 1985)، وتنقسم نماذج نظرية الاستجابة للفقرة إلى نماذج أحادية البُعد، وهي التي تحدّد لكلّ شخص سمة واحدة، حيث تُعَدّ هذه النماذج مناسبة للفقرات التي تتضمّن سمة واحدة أو مجموعة من السمات الثابتة عبر الفقرات، ونماذج متعدّدة الأبعاد، حيث تُعَدّ هذه النماذج مناسبة للبيانات التي تُسْهم فيها أكثر من سمة في الإجابة عن الفقرات، وتنقسم كلّ من النماذج أحادية البُعد والنماذج متعددة الأبعاد إلى قسمين: القسم الأول: نماذج ثنائية الاستجابة، وفيها تكون استجابة الفرد إما إجابة صحيحة؛ فيأخذ الدرجة 1 أو إجابة خاطئة؛ فيأخذ الدرجة 0، وأكثر هذه النماذج شيوعًا هي النماذج اللوجستية نظرًا لشكلها البسيط حسابيًا، ومنها: النموذج اللوجستي أحادي المعلمة، والنموذج اللوجستي ثنائي المعلمة، والنموذج اللوجستي ثلاثي المعلمة (Embretson and Raise, 2000)، والقسم الثاني: نماذج متعددة الاستجابة، وفيها تكون استجابة الفرد متدرّجة مثل حلّ المسائل الرياضية، فقد تكون درجة

الفرد غير كاملة، فتكون درجته مثلًا الدرجة 3 من 5 (Reckase, 2009)، وسنتناول في هذه الدراسة النماذج أحادية البعد بقسميها.

وبالرجوع إلى De Ayala (2009)، و(2009) Reckase، و Hambelton and Swaminathan (1985)، و(2014) Kolen and Brennan، و Ogunsakini and Shogbesan (2018)، و(2000) Embretson and Raise فإننا نستعرض بعضًا من أشهر النماذج أحادية البُعد.

1. النماذج ثنائية الاستجابة:

- **النموذج اللوجستي أحادي المعلمة One Parameter Logistic Model (1PLM):** يُسَمّى هذا النموذج بنموذج راش، وهو أحد نماذج نظرية الاستجابة للفقرة ثنائية الاستجابة، ويُعَدّ النموذج الأكثر استعمالًا من بين نماذج نظرية الاستجابة للفقرة؛ وذلك لسهولة التعامل معه مقارنة بالنماذج الأخرى، ويفترض هذا النموذج أن فقرات الاختبار تختلف في صعوبتها فقط، ويُعَبّر عن هذا النموذج بدالة رياضية تربط بين احتمال إجابة الطالب j عن الفقرة إجابة صحيحة P_i وقدرة الطالب θ وصعوبة الفقرة b_i :

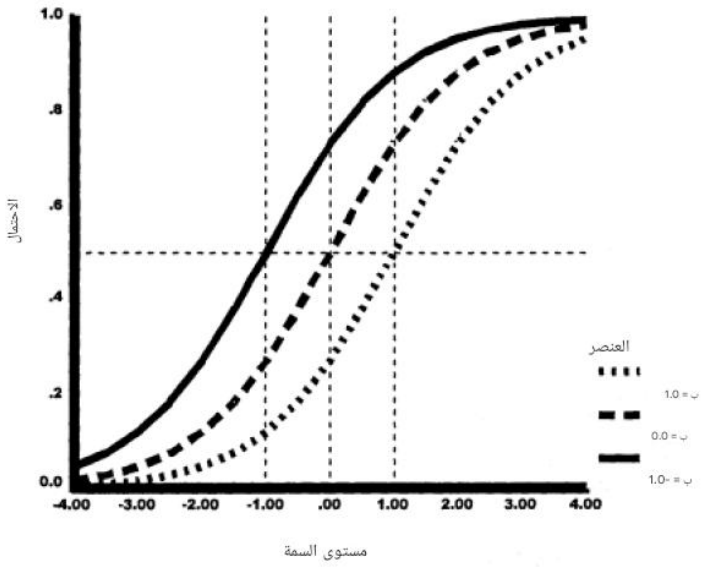
$$P_i(\theta_j) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}}$$

ويوضح الشكل (1) منحنيات خصائص الفقرة ICC لثلاث فقرات من النموذج أحادي المعلمة، حيث تمتاز منحنيات خصائص الفقرة في النموذج أحادي المعلمة بـ:

- تزداد احتمالات تدريجيًا مع زيادة مستوى السمة (القدرة).
- تتقارب المنحنيات، لكنها لا تتقاطع.

– يتحوّل معدل التغير في الاحتمال من زيادات متسارعة إلى زيادات متباطئة عند نقطة انعطاف منحنيات خصائص الفقرة، ويحدث ذلك عندما يصل احتمال الإجابة عن الفقرة إلى 0.50، وتشير الخطوط المرجعية في الشكل (1) إلى أنه عند احتمال الإجابة عن الفقرة 0.50 يشير مستوى السمة إلى مستوى صعوبة الفقرة (Embretson and Raise, 2000).

الشكل (1) منحنيات خصائص الفقرة في النموذج أحادي المعلمة



- **النموذج اللوجستي ثنائي المعلمة (Tow-Parameter Logistic Model (2PLM**:

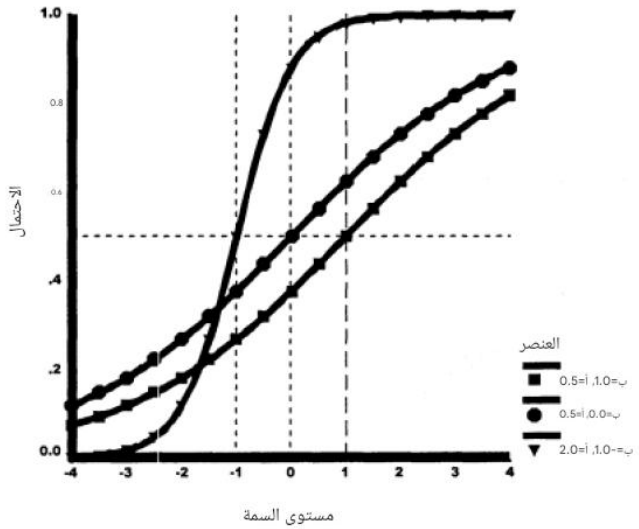
يُعَدّ هذا النموذج أكثر تعقيدًا من النموذج السابق، ويفترض هذا النموذج أن فقرات الاختبار تختلف في صعوبتها، وقدرتها التمييزية، وهو عبارة عن دالة رياضية تربط بين احتمال إجابة الطالب عن الفقرة إجابة صحيحة P_i، وقدرة الطالب θ، وصعوبة الفقرة b_i، وتمييز الفقرة a_i:

$$P_i(\theta_j) = \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}}$$

حيث يُعَدّ هذا النموذج مناسبًا للمقاييس التي لا ترتبط فيها الفقرات بشكل متساوٍ بالسمة الكامنة (بمعنى لا تشير الفقرات بشكل متساوٍ إلى مقدار امتلاك الشخص للسمة الكامنة)، ويوضح الشكل (2) منحنيات

خصائص الفقرة لثلاث فقرات في النموذج ثنائي المعلمة، حيث تختلف الفقرات الثلاث في صعوبتها وتمييزها، ونلاحظ من الشكل (2) أن منحنى الفقرة 2 يتقاطع مع منحنيات الفقرتين: 1 و3، ونلاحظ أن مستوى السمة عند احتمال 0.50 يشير إلى صعوبة الفقرة (Embretson and Raise, 2000).

الشكل (2) منحنى خصائص الفقرة في النموذج ثنائي المعلمة



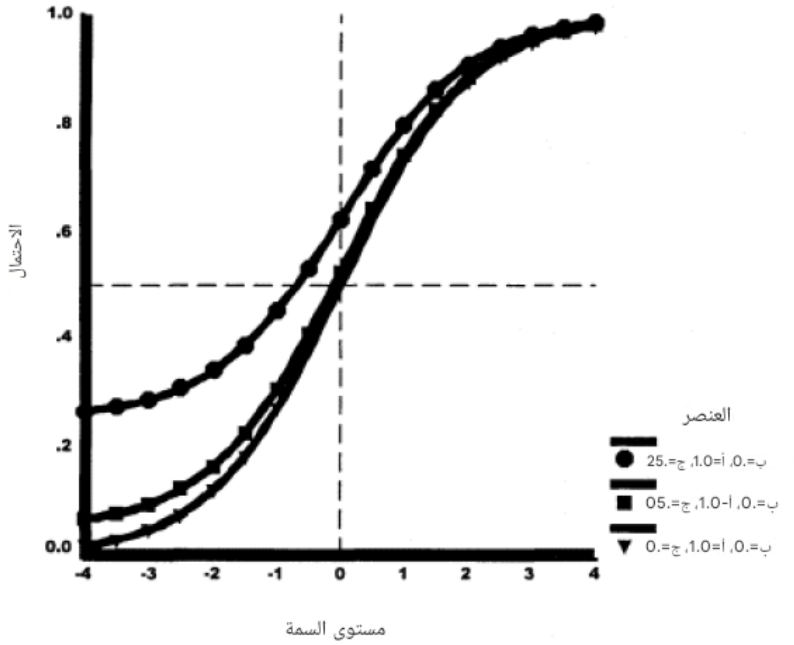
- **النموذج اللوجستي ثلاثي المعلمة Three-Parameter Logistic Model (3PLM):** يراعي هذا النموذج إمكانية اختلاف فقرات الاختبار في صعوبتها، وقوتها التمييزية، واحتمالية حلّها عن طريق التخمين؛ ولذا يكون احتمال الإجابة عن الفقرة أكبر عن النصف حتى عند مستويات القدرة المنخفضة، فهذا النموذج عبارة عن دالة رياضية تربط بين احتمال إجابة الطالب عن الفقرة إجابة صحيحة P_i، وقدرة الطالب θ، وصعوبة الفقرة b_i، وتمييز الفقرة a_i واحتمالية الإجابة عن الفقرة باستعمال التخمين c_i:

$$P_i(\theta_j) = c_i + (1 - c_i)\ \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}}$$

ويوضح الشكل (3) منحنيات خصائص الفقرة لثلاث فقرات متساوية في الصعوبة والتمييز، ولكن مع خطوط تقارب سفلية مختلفة؛ ولذا يكون احتمال الإجابة عن الفقرة أكبر من الصفر حتى عند

المستويات الدنيا من السمة. وتجدر الإشارة إلى أن صعوبة الفقرة لها تفسير مختلف في النموذج ثلاثي المعالم على الرغم من أنها لا تزال تحدث عند نقطة الانعطاف في منحنى خصائص الفقرة، إلا أنها لم تَعِد تمثّل مستوى السمة الذي يكون عنده احتمال الإجابة 0.50، حيث يظهر من الشكل (3) ثلاث فقرات بمستوى الصعوبة نفسه وهو 0.00، حيث إن صعوبة الفقرة مرتبطة باحتمالية أكبر بكثير من 0.50 للفقرة ذات التخمين البالغ 0.25 (Embretson and Raise, 2000).

الشكل (3) منحنيات خصائص الفقرة في النموذج ثلاثي المعلمة



2. نماذج تحتوي على أكثر من فئتين من الدرجات:

تدرس هذه النماذج خصائص استجابات المفحوصين على أنواع مختلفة من أنواع فقرات الاختبار، والتي تحتوي على أكثر من فئتين من الدرجات مثل فقرات الكتابة، وفقرات حلّ المسائل الرياضية، وفقرات الاستبانة متعددة التدريج، ومن هذه النماذج:

- **نموذج الاستجابة المتدرجة Graded Response Model (GRM)**: يُعَدّ هذا النموذج تعميمًا للنموذج ثنائي المعلمة (2PLM)، حيث يفترض هذا النموذج أن الإجابة عن فقرات الاختبار تتطلّب عددًا من الخطوات؛ فالإنجاز الصحيح للخطوة يتطلّب الإنجاز الصحيح للخطوات السابقة، وكذلك فإن هذا النموذج لا يتطلب تساوي عدد فئات الإجابات لكلّ الفقرات، ويشير التقي (2008) إلى أن هذا النموذج يستعمل في حالة عدم تساوي مستويات التقدير للفقرات، وفي هذا النموذج فإن كل فقرة يمكن وصفها بمعلم تمييز واحد a_i ، والذي يقع ما بين معالم الفئات أو الفواصل لهذه الفئات، وهناك مرحلتان في عملية حساب احتمالية الإجابة لكلّ فئة، ولفهم هذه المراحل نفترض أنا لدينا فقرة من 4 درجات، فيحصل المفحوص على درجات بين 0 – 4؛ ولذا يكون عدد الفواصل 4، وهي الحدود بين الدرجات $m_j = 4$ و$j = 1,2,3,4$ (Reckase, 2009).

وتمثل الخطوة الأولى في هذا النموذج تقدير احتمالية الإجابة عن الفقرة، والتي تعتمد على عملية تقدير منحنيات الفئة m_i لكلّ فقرة باستعمال المعادلة الآتية:

$$P_{ix}(\theta) = \frac{e^{a_i(\theta - b_{ij})}}{1 + e^{a_i(\theta - b_{ij})}}$$

حيث إن x= j = 1,2,……,m

$P_{ix}(\theta)$: احتمال الإجابة في الفئة j.

a_i : تمييز الفقرة.

b_{ij} : معامل صعوبة الفقرة.

θ : معلمة القدرة.

عند التدقيق في المعادلة السابقة فنلاحظ أنها تمثل النموذج ثنائي المعلمة من نماذج نظرية الاستجابة للفقرة، حيث إنها تمثل احتمال حصول المفحوص على الدرجة الخام x التي تقع في الحدّ الفاصل للفئة أو أعلى منه j = 1,2,……,m (Reckase, 2009).

وذكر Embretson and Raise (2000) أنه حسب النموذج المتدرج يجب تقدير منحنى واحد لكلّ حدّ أو فاصل بين الفئات، وبذلك فإن الإجابة المكونة من 5 فئات لها 4 معالم صعوبة، ومعلم تمييز واحد، ويذكر التقي (2008) أن صعوبة الفواصل في هذا النموذج متزايدة، بمعنى أن صعوبة الفاصل الثاني أكبر من صعوبة الفاصل الأول، وبشكل عام فإن في هذا النموذج يكون التعامل مع الفقرات على أنها سلسلة من الاستجابات $m_i = k - 1$ ، حيث k تمثل عدد الفئات، ويكون تقدير النموذج الثنائي لكلّ سلسلة مع التركيز على أن معلم التمييز ثابت، وتُحْتَسَب احتمالية الاستجابة لكلّ فئة بواسطة المعادلة الآتية:

$$P_{ix}(\theta) = P^*_{ix}(\theta) - P^*_{i(x+1)}(\theta)$$

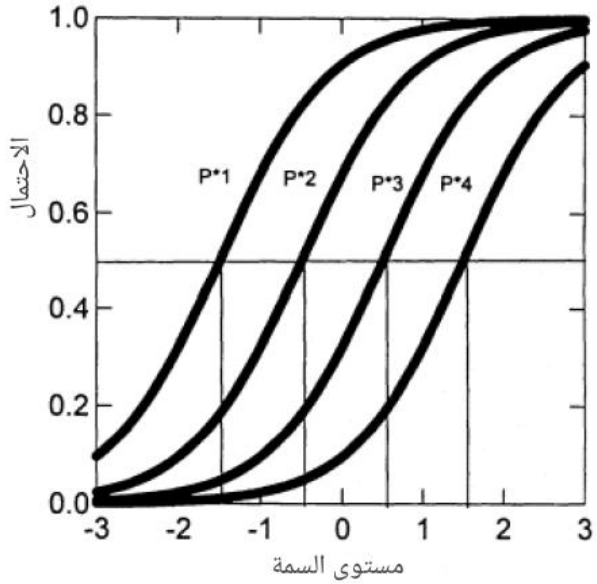
حيث، $P_{ix}(\theta)$: هو احتمال حصول الفرد الذي قدرته θ على الدرجة x للفقرة i.

وحسب التعريف فإن احتمال الإجابة في أدنى فئة أو أعلى منها هو $P^*_{i0}(\theta) = 1$ ، وأن احتمال الإجابة فوق أعلى فئة هو $P^*_{i5}(\theta) = 0$ ؛ ولذا يكون احتمال الإجابة في كلّ فئة من الفئات الخمس كما يأتي:

$$P^*_{i0}(\theta) = 1 - P^*_{i1}(\theta)$$
$$P^*_{i1}(\theta) = P^*_{i1} - P^*_{i2}(\theta)$$
$$P^*_{i2}(\theta) = P^*_{i2} - P^*_{i3}(\theta)$$
$$P^*_{i3}(\theta) = P^*_{i3} - P^*_{i4}(\theta)$$
$$P^*_{i4}(\theta) = P^*_{i4} - 0$$

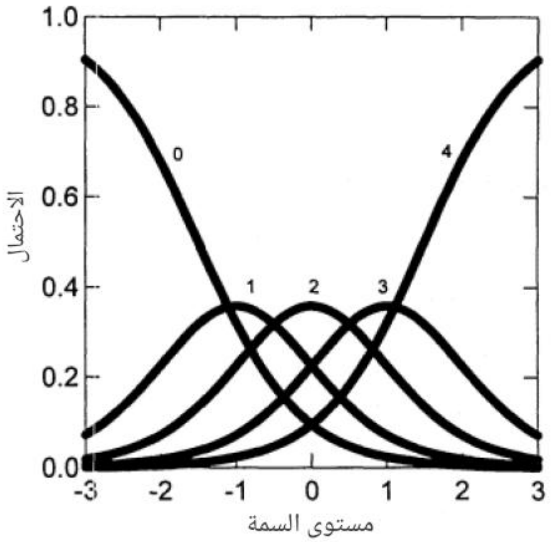
ويوضح الشكل (4) منحنيات خصائص الفئات لفقرة تتكوّن من خمس فئات استجابة، والتي تمثل الفرق بين منحنيات تراكمية وليس احتمال الفئة نفسها، حيث إن: $a_i = 1.5$، $b_{i1} = -1.5$، $b_{i2} = -0.5$، $b_{i3} = 0.5$، $b_{i4} = 1.5$.

الشكل (4) منحنيات خصائص الفئات لفقرة تتكون من خمس فئات وفق نموذج GRM



ويوضح الشكل (5) منحنيات الخصائص التشغيلية للفقرة نفسها، حيث توضح هذه المنحنيات احتمال حصول الفرد على فئة معينة (احتمال الفئة)، وذلك حسب قدرة الفرد، حيث نلاحظ أنه لأي قيمة لـ θ فإن مجموع احتمالات الاستجابة يساوي 1.

الشكل (5) منحنيات خصائص استجابة الفئات لفقرة ذات خمس فئات استجابة وفق نموذج GRM



- **نموذج الاستجابة المتدرجة المُعدّل (M– Modified Graded Response Model (GRM)** طوّر Muraki تعديلًا على نموذج الاستجابة المتدرجة؛ ليسهل استعماله في تحليل صيغ الاستبانات ذات المقياس نفسه (تحتوي جميع الفقرات على عدد فئات الاستجابة نفسها)، مثل مقياس ليكرت، ويسمح هذا النموذج لمعلمة التمييز بالاختلاف بين الفقرات، ويُعَدّ هذا النموذج غير مباشر، حيث يتطلب حساب احتمال الاستجابة القيام بخطوتين، حيث يقسّم هذا النموذج عتبة الصعوبة إلى حدّين هما معلمة الموقع b_i، ومجموعة معلمات عتبة الفئة c_j للمقياس بأكمله، أي أن $b_{ij} = b_i c_i$.

والفرق بين نموذج الاستجابة المتدرجة (GRM) ونموذج الاستجابة المتدرجة المُعدّل (M–GRM) هو أنه في نموذج (GRM) تقدّر مجموعة واحدة من معلمات عتبة الفئة b_{ij} لكلّ فقرة من فقرات المقياس، بينما في نموذج (M–GRM) تقدّر مجموعة واحدة من معلمات عتبة الفئة c_j للمقياس بأكمله، ويُقدّر معلم موقع واحد لكلّ فقرة b_i؛ حيث تشير معلمة موقع الفقرة إلى صعوبة الفقرة أو قيمة المقياس.

ويمكن كتابة منحنيات الخصائص التشغيلية لهذا النموذج على الصورة:

$$P_{ix}^{*}(\theta) = \frac{e^{a_i(\theta-(b_i-c_i))}}{1+e^{a_i(\theta-(b_i-c_i))}}$$

ولذا فإن احتمال الاستجابة في فئة معينة يساوي:

$$P_{ix}(\theta) = P_{ix}^{*}(\theta) - P_{i(x+1)}^{*}(\theta)$$

- **نموذج التقدير الجزئي Partial Credit Model (PCM)**: طُوِّر هذا النموذج لتحليل فقرات الامتحان التي تتطلّب الإجابة عنها خطوات متعددة، حيث يمكن استعماله في الاختبارات التحصيلية مثل اختبارات الرياضيات، حيث من الممكن أن يحصل الطالب على جزء من الدرجة، وتُعَبِّر درجة الفقرة عن عدد الخطوات التي أُكْمِلت بنجاح، فكلما أنجز الممتحن جزءًا أكبر من المهمة حصل على درجة أعلى، ويختلف هذا النموذج عن نموذج الاستجابة المتدرجة كونه نموذجًا مباشرًا لنظرية الاستجابة للفقرة، حيث يمكن كتابته في صورة حاصل قسمة دالتين أسيتين، وكذلك يمكن حسبان هذا

النموذج امتدادًا للنموذج اللوجستي أحادي المعلمة، حيث يشتمل على جميع خصائص نموذج راش مثل الفصل بين معالم الفقرات وقدرات الأفراد (Embretson & Reise, 2000).

وعلى افتراض أن لفقرة درجات $X = 0, \ldots\ldots, m_i$ ، فيكون لها فئات استجابة عددها $k_i = m_i + 1$ ؛ ولذا يعبّر عن هذا النموذج رياضيًا بالعلاقة:

$$P_{ix}(\theta) = \frac{e^{\sum_{j=0}^{x}(\theta - \delta_{ij})}}{\sum_{\tau=0}^{m_i}\left[e^{\sum_{j=0}^{\tau}(\theta - \delta_{ij})}\right]}$$

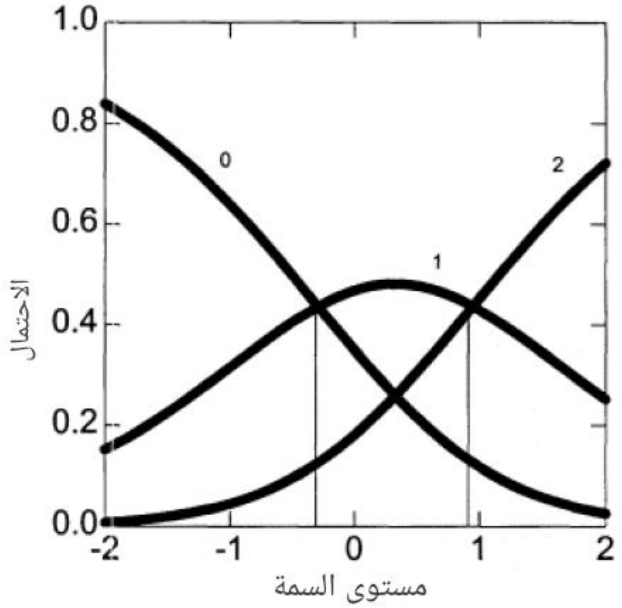
$$\sum_{j=0}^{0}(\theta - \delta_{ij}) = 0 \qquad \text{حيث}$$

حيث δ_{ij} تمثل صعوبة الخطوة (j) ، وحيث $(\mathrm{j} = 1, \ldots\ldots, m_i)$ من الفقرة i ، فكلما زادت قيمة δ_{ij} زادت صعوبة الخطوة داخل الفقرة مقارنة بالخطوات الأخرى للفقرة نفسها؛ فقد تكون بعض الخطوات داخل الفقرة أصعب أو أسهل من غيرها.

كذلك فإن δ_{ij} تمثل النقطة على مقياس السمة التي يتقاطع عندها منحنيان متتاليان لاستجابة الفئة، ولذلك تُسَمّى δ_{ij} في نموذج التقدير الجزئي (PCM) باسم تقاطعات الفئات، وكذلك يفترض هذا النموذج أن معلم التمييز ثابت لجميع الفقرات، حيث يُعَدّ امتدادًا لنموذج راش.

ويوضح الشكل (6) منحنيات استجابة الفئات لفقرة مكونة من 3 فئات، فبالنسبة لهذه الفقرة تكون معلمات δ_{ij} هي $\delta_{i1} = -0.29$، و$\delta_{i2} = 0.95$ ، حيث تمثل هذه النقاط نقاط تقاطع منحنيات استجابة الفئات، ويمكن استعمال منحنيات استجابة الفئات لحساب الاستجابة الأكثر احتمالية للممتحن عند نقاط مختلفة من السمة (القدرة).

الشكل (6) منحنيات استجابة الفئات لفقرة مكونة من ثلاث فئات وفق نموذج PCM



وعلى سبيل المثال ومن الشكل (6) فإن الاستجابة الأكثر احتمالية لممتحن عند مستوى قدرة $\theta = -2$ هي الدرجة صفر، بينما الاستجابة الأكثر احتمالية لممتحن عند مستوى قدرة $\theta = 1.5$ هي الدرجة 2.

- **نموذج التقدير الجزئي المعمم Generalized Partial Credit Model (G-PCM):** قام (Muraki,1993) بتطوير نموذج التقدير الجزئي، بحيث يسمح للفقرات في اختبار ما بالاختلاف في معلم التمييز، أي أن هذا النموذج يسمح بالتباين في القدرة التمييزية للفقرات، فبدلًا من أن يستعمل موراكي النموذج أحادي المعلمة (نموذج راش) استعمل النموذج ثنائي المعالم، ويُعَدّ هذا النموذج مناسبًا للفقرات المفتوحة التي يكون تقييمها وفق مجموعة من الخطوات أو المعايير مثل مسائل الرياضيات، حيث إن درجة الفقرة في الاختبار تُعبّر عن عدد الخطوات (المعايير) التي أُكْمِلَت بنجاح، حيث تعني الدرجة الأعلى أن الممتحن قد أنجز جزءًا أكبر من المهمة المطلوبة.

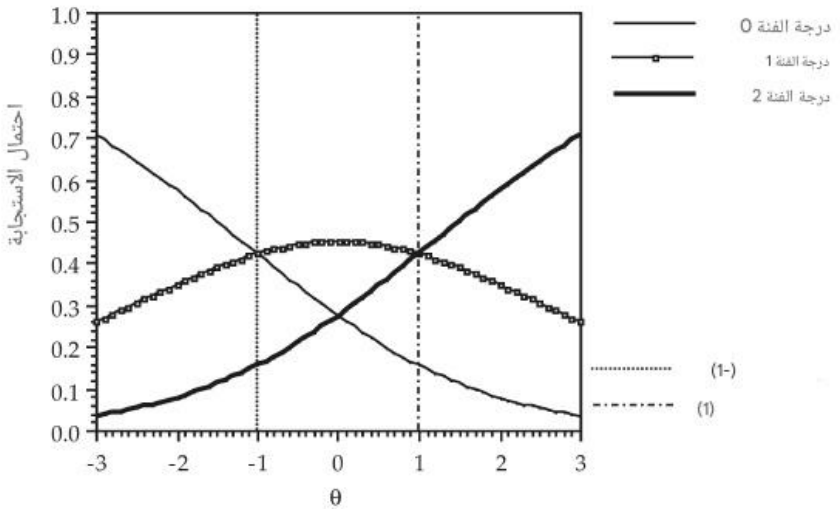
ويمكن التعبير عن هذا النموذج بالعلاقة الرياضية التالية (Embretson & Reise, 2000):

$$P_{ix}(\theta) = \frac{e^{\sum_{j=0}^{x} a_i(\theta-\delta_{ij})}}{\sum_{\tau=0}^{m_i}\left[e^{\sum_{j=0}^{\tau} a_i(\theta-\delta_{ij})}\right]}$$

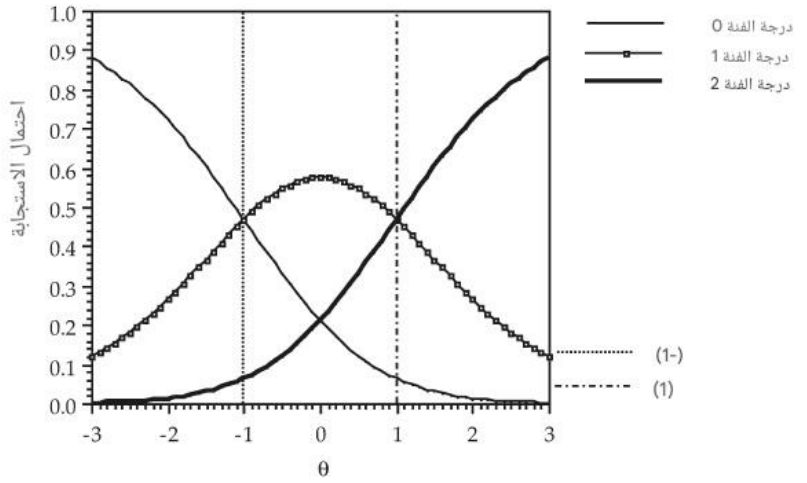
حيث $\sum_{j=0}^{0}(\theta-\delta_{ij})=0$ ، a_i تمثل تمييز الفقرة، و δ_{ij} تمثل النقطة على مقياس السمة التي يتقاطع عندها منحنيان متتاليان لاستجابة الفئة؛ ولذلك تُسَمَّى δ_{ij} بالتسمية نفسها في نموذج التقدير الجزئي(PCM) باسم تقاطعات الفئات، وهي النقطة التي تكون عندها احتمالية الاستجابة في فئة معينة تساوي احتمالية الإجابة في الفئة السابقة، ويعتمد تفسير تمييز الفقرات في النماذج متعددة الاستجابة على معامل التمييز وانتشار عتبات الفئات (مثلًا: في نموذج GRM أو M–GRM)، أو يعتمد على معامل التمييز وتقاطعات الفئات (مثلًا: في نموذج PCM أو G–PCM).

وتوضح الأشكال: (7)، و(8)، و(9)، و(10) منحنيات استجابة الفئات لأربع فقرات ذات 3 فئات استجابة مع اختلافها في التمييز حيث: $a_1=0.5$، $a_2=1$، $a_3=2$، $a_4=2$ على التوالي وفق نموذج G–PCM.

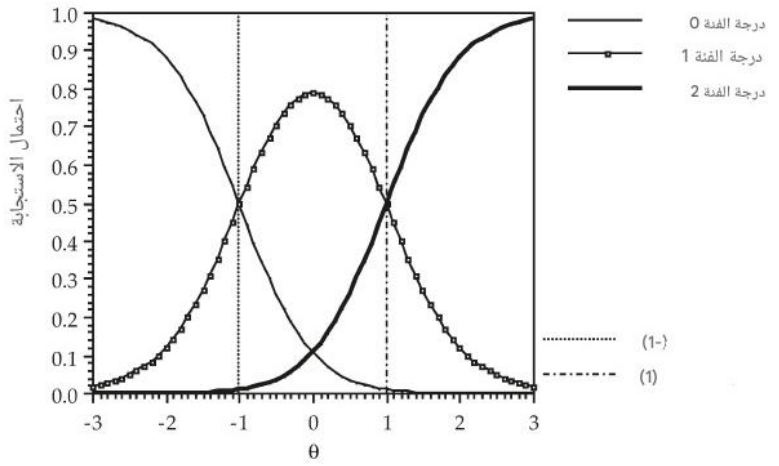
الشكل (7) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a_1=0.5$



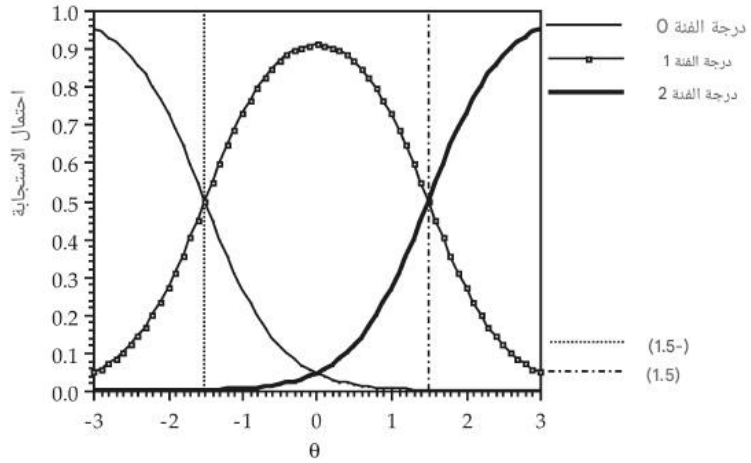
الشكل (8) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a_2 = 1$



الشكل (9) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a_3 = 2$



الشكل (10) منحنى استجابة الفئة لفقرة ذات 3 فئات استجابة بتمييز $a_3 = 2$



حيث توضح هذه المنحنيات نقاط تقاطعات الفئات، فمثلًا في الأشكال: (7)، و(8)، و(9) تكون نقاط تقاطعات الفئات عند $\theta = -1$، $\theta = 1$، بينما في الشكل (10) نقاط تقاطعات الفئات $\theta = -1.5$، $\theta = 1.5$، وكذلك تظهر هذه الأشكال تأثير اختلاف التمييز في شكل المنحنيات.

- **نموذج مقياس التقدير Rating Scale Model (RSM)**: يُعَدّ نموذج مقياس التقدير (RSM) الذي طوره أندريش أحد النماذج المناسبة لبيانات الاستجابة في مقياس ليكرت، حيث يستعمل نموذج (RSM) استجابات تأتي من سلسلة من الفئات الترتيبية، ويفترض هذا النموذج أن هذه الفئات الترتيبية يفصلها عن بعضها سلسلة من العتبات الترتيبية، وموقع كلّ عتبة على متصل السمة يمثّل الحدّ الفاصل بين فئتي استجابة متجاورتين، حيث إن عدد العتبات m أقل بواحد من عدد فئات الاستجابة (عدد فئات الاستجابة = m +1)، وهذا مختلف عن نموذج التقدير الجزئي؛ حيث إنه في نموذج التقدير الجزئي يمكن للفقرات أن تختلف في عدد العتبات وفي عدد فئات الاستجابة، بينما في هذا النموذج (RSM) يجب أن يكون لجميع الفقرات العدد نفسه من فئات الاستجابة والعدد نفسه من العتبات، وتأخذ العتبات القيم نفسها في جميع الفقرات، ولكن ليس بالضرورة أن تأخذ المواقع نفسها على متصل السمة لجميع الفقرات، ويتفق هذا النموذج مع نموذج (PCM) في تساوي جميع الفقرات في معلمة التمييز.

ففي هذا النموذج تُقْسَم نقاط تقاطعات الفئات δ_{ij} إلى قسمين هما δ_j، λ_i، حيث δ_j هي معلمات تقاطع الفئات، و λ_i هي موقع الفقرة على مقياس السمة.

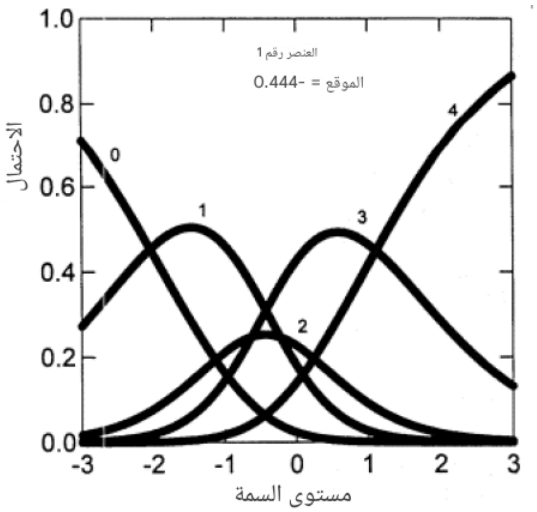
وقد صاغ Dodd *نموذج* (RSM) *في الصيغة الرياضية الآتية:*

$$P_x(\theta) = \frac{e^{\sum_{j=0}^{x}(\theta-(\lambda_i+\delta_j))}}{\sum_{\tau=0}^{m}\left[e^{\sum_{j=0}^{x}(\theta-(\lambda_i+\delta_j))}\right]}$$

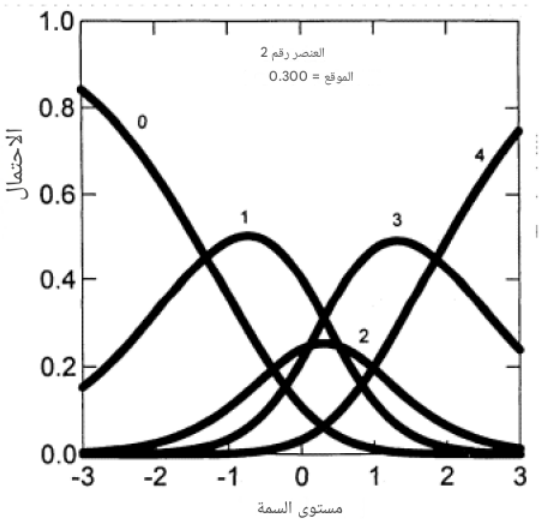
حيث $\sum_{j=0}^{0}(\theta-(\lambda_i+\delta_j)=0$

ويوضح الشكلان: (11، و12) منحنيات استجابة الفئات لفقرتين ذات أربع فئات استجابة وفق نموذج (RSM).

الشكل (11) منحنى استجابة الفئة للفقرة الأولى وفق نموذج (RSM) حيث $\lambda_i = -0.444$



الشكل (12) منحنى استجابة الفئة للفقرة الثانية وفق نموذج (RSM) حيث $\lambda_i = 0.333$



نلاحظ من الشكلين: (11)، و(12) أن الشكل العام للمنحنيات هو نفسه، لكن منحنيات الفقرة الثانية مُزاحة لليمين بالنسبة للفقرة الأولى، حيث يمثل معامل الموقع λ_i متوسط صعوبة الفقرة بالنسبة لتقاطعات الفئات، وكما هو الحال في نموذج التقدير الجزئي يكون مجموع الاحتمالات لجميع الفئات عند مستوى قدرة معينة يساوي 1.

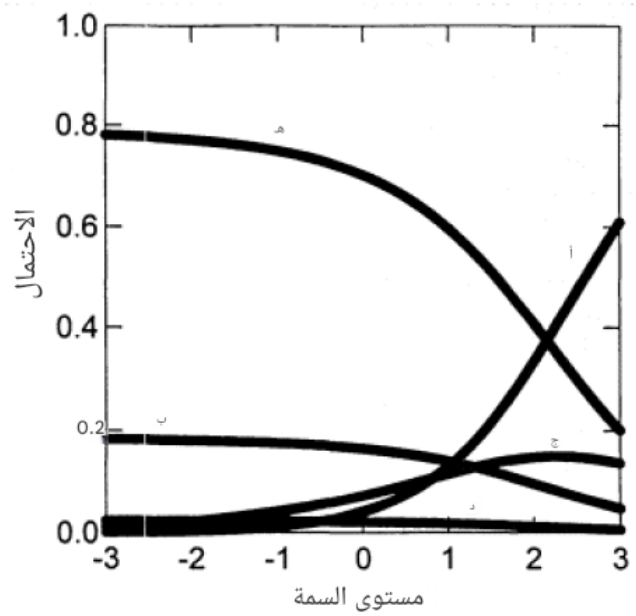
- **نموذج الاستجابة الاسمية Nominal Response Model (NRM)**: قدم Bock نموذجًا يمكن استعماله لتحليل الاستجابات لفقرات ذات فئات استجابة غير مرتبة على متصل السمة مثل أداة لجمع أحكام المقيمين؛ حيث إن أحكام المقيمين لا تكون وفق تصـــنيفات مرتبة. وفي نموذج (NRM) يمكن كتابة احتمال استجابة الممتحن في الفئة x حيث $(x = 0, 1, \dots\dots\dots, m_i)$ على الصورة:

$$P_{ix}(\theta) = \frac{e^{(a_{ix}\theta + c_{ix})}}{\sum_{x=0}^{m} e^{(a_{ix}\theta + c_{ix})}}$$

حيث a_{ix} تمثل التمييز للفئة x للفقرة i ، و c_{ix} تمثل معلمة الاعتراض الفئة x (المقطع الصادي)، وكذلك فإن $\sum a_{ix} = \sum c_{ix} = 0$، وفي هذا النموذج معلمة a_{ix}، ومعلمة c_{ix} يجب تقديرها لكلّ فئة من فئات الاستجابة $(m_i + 1)$.

ويوضح الشكل (13) منحنيات استجابة الفئة لفقرة من خمس فئات استجابة وفق نموذج الاستجابة الاسمية.

الشكل (13) منحنيات استجابة الفئة لفقرة من خمس فئات استجابة وفق (NRM)



وفي هذه الدراسة استعمل الباحث النموذج ثلاثي المعالم (3PLM) لمعايرة الفقرات الثنائية ونموذج التقدير الجزئي المعمم لمعايرة الفقرات المقالية (متعددة التصحيح)؛ وذلك لملائمتهما لطبيعة بيانات الدراسة، حيث إن الفقرات ثنائية التصحيح في صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام تختلف في قدرتها التمييزية، وكذلك يكون التخمين فيها ممكنًا، وكذلك فإن نموذج التقدير الجزئي المعمم مناسب لتحليل اختبارات التحصيل مثل الرياضيات، حيث تكون الإجابة الصحيحة جزئيًا ممكنة (Embretson & Reise, 2000)، ويُعَدّ نموذج التقدير الجزئي المعمم مناسبًا للفقرات المفتوحة، وذلك عندما يأخذ المصححون في الحسبان الكثير من مكونات الاستجابة المختلفة، ويجعلون لكلّ منها صحيح أو غير صحيح، وتكون درجة الفقرة عدد الخطوات التي أُكْمِلَت بنجاح، وتمثل الدرجات على الفقرة مستويات الأداء، حيث تعني الدرجة الأعلى إنجاز خطوات أكثر (Reckase, 2009)، ويفترض هذا النموذج أن الصعوبة غير متدرجة، فقد تكون صعوبة الفاصل الأول أعلى من صعوبة الفاصل الثاني وهكذا (التقي، 2008).

كما أن نموذج التقدير الجزئي المعمم (G–PCM) يستعمل في تقدير معالم فقرات اختبارات TIMSS وPIRLS (Bezirhan & von Davier, 2023; Yamamoto & Kulick, 2000).

معادلة درجات الامتحانات Test Score Equating

تشكّل الاختبارات إحدى أهم أدوات القياس النفسيّ والتربويّ؛ حيث إنها تزوّد متخذي القرار بالمعلومات اللازمة والضرورية لتصنيف الطلبة إلى فئات مختلفة وفقًا لأدائهم في الاختبارات (الصمادي وعدس، 2008)، ونظرًا لأهمية القرارات التربوية المستمدة من الاختبارات، وللحفاظ على سرية الاختبارات، ولتجنب إعادة تطبيق الاختبار نفسه على مجموعة الطلبة نفسها أو مجموعة أخرى، فإنه يُلْجَأ إلى إعداد عدة نماذج أو صور للاختبار، ولكن قد تختلف الصور المختلفة للاختبار في صعوبتها، وكذلك تختلف مجموعة الطلبة الذين ستُطَبَّق عليهم هذه الصور، وحتى يمكن تعديل الاختلاف في

الصعوبة يُلْجَأ إلى أساليب إحصائية لمعادلة الدرجات، بحيث تستعمل الدرجات على صور الاختبار بشكل تبادليّ (Kolen & Brennan, 2014)، وتكون المعادلة ناجحة إذا كان للطلبة الدرجات المتوقعة نفسها بغضّ النظر عن أيّ صورة تقدموا لها، فالمعادلة تعمل على تحقيق عدالة الاختبار ومقارنة أداء الطلبة عبر الصور المختلفة (المحروق وعليان، 2011).

ويقصد بمعادلة صورتي الاختبار أنها عملية تحويل الدرجات على صور الاختبار إلى مقياس مشترك موحّد، بحيث تصبح القياسات المستمدة من درجات كلّ من الصورتين متكافئة بعد إجراء هذا التحويل، وتهدف المعادلة إلى تعديل الفروقات في الصعوبة، بحيث يمكن إجراء مقارنات بين الطلبة بشكل متبادل (العنزي، 2015). وتُعرَّف أيضًا بأنها عملية الحصول على الدرجات المتكافئة والمتساوية لأداتين تقيسان السمة نفسها بمقدار متساوٍ من الثبات مع تساوي الرتب المئينية المتناظرة للدرجات (Croocker & Algina, 1986). ويُعرِّفها الدوسري (2001) بأنها إجراء لإزالة تأثير الفرق بين متوسط مستويات الصعوبة لفقرات الاختبار وكذلك الفرق بين متوسط التمييز لفقرات الاختبار على الدرجات، بحيث تصبح صورتا الاختبار متكافئتين. وتُعرِف الزعابي وآخرون (2018) عملية تعادل صورتي الاختبار بأنها عملية تحويل وحدات نظام القياس لإحدى الصورتين إلى وحدات نظام قياس الصورة الأخرى، بحيث تصبحان متساويتين في القياس بعد التحويل.

ويَرى (Lord, 1982) أن صور الاختبار تكون متكافئة عندما لا يؤدي تطبيق أيّ صورة منها إلى فرق أو اختلاف في أداء الطالب، حيث لا يمكن حسبان أن صورتي الاختبار متكافئتان إلا إذا كانت القدرة المستنتجة للطالب نفسها بغضّ النظر عن الصورة التي قُدِّمَت للطالب، وقد اتفق معه في هذا التعريف الحربي وعليان (2009)، حيث عرّفاها بأنها إجراء إحصائي يعمل على وضع درجات

الصور المختلفة للاختبار على تدريج مشترك، وبذلك يمكن الحصول على تقدير قدرة الفرد نفسها بغضّ النظر عن الصورة الاختبارية المقدمة له، وهناك نوعان من أنواع معادلة الاختبارات هما:

1- المعادلة الأفقية Horizontal Equating:

تهدف المعادلة الأفقية إلى تعديل الفروق الناتجة عن اختلاف مستوى الصعوبة، وذلك عندما تكون الصور المختلفة متقاربة في صعوبتها، وهذا يعني أن المعادلة الأفقية تكون مناسبة عندما يكون لدينا صور مختلفة ومتوازية إحصائيًا ورياضيًا من الاختبار نفسه، ولكنها مختلفة في الفقرات (الدوسريّ، 2001)، والمثال عليها هو معادلة الصور المختلفة للاختبار عبر سنوات مختلفة، حيث تُعَدّ المعادلة الأفقية مناسبة عندما تكون هناك حاجة إلى صور مختلفة للاختبار لأسباب تعود إلى المحافظة على سرية الاختبار أو أسباب أخرى؛ فالصور المختلفة لن تكون متطابقة، ولكنها تكون متوازية، وكذلك فإن توزيعات قدرات الممتحنين الذين تطبق عليهم هذه الصور متساوية (Hambelton & Swaminathan, 1985).

2- المعادلة الرأسية Vertical Equating:

وفيها تكون صور الاختبار مختلفة في مستوى صعوبتها، ولن تقيس النطاق نفسه من المحتوى أو المهارات وكذلك مجموعات الأفراد (الطلبة) الذين تطبق عليهم هذه الصور مختلفة في مستويات القدرة، حيث تهدف المعادلة الرأسية إلى تحويل درجات الاختبارات المصممة لمجتمعات من الأفراد في مستويات تعليمية مختلفة إلى تدريج واحد، حيث تكون عندما تُنْجَز مقارنات بين مجموعات من الطلبة من مستويات صفية مختلفة، وكذلك عندما نريد إجراء معادلة درجات بطاريات مختلفة من الاختبارات المصممة لصفوف مختلفة، وينبغي ملاحظة أن نتائج كلّ من المعادلة الأفقية والرأسية هو وضع درجات صور مختلفة للاختبار على التدريج نفسه (علي وآخرون، 2016).

وتتيح المعادلة الرأسية مقارنة قدرات الممتحنين في مستويات مختلفة؛ ولذا فلن تكون هناك صور مختلفة للاختبار نفسه، بل ستكون هناك صور اختبارية بمستويات صعوبة مختلفة، وكذلك فإن توزيعات قدرات الممتحنين ستكون مختلفة (Hambelton & Swaminathan, 1985).

ويشير الحربيّ وعليان (2009) إلى أن الفارق بين نوعي المعادلة يعتمد على مدى الاختلاف في صعوبة الصور المراد معادلتها، بالإضافة إلى انتماء المفحوصين إلى صف دراسي واحد (أي تقيس مستوى القدرة نفسها) أو انتمائهم إلى صفوف مختلفة (أي تقيس مستويات مختلفة من القدرة).

شروط (متطلبات) معادلة درجات الاختبارات Requirements of Equating :

تُعَدّ هذه المتطلبات مبادئ أساسية عند إجراء عملية المعادلة؛ حيث تنطلق هذه المتطلبات من منطلق أنه يجب أن يحصل الطالب على الدرجات نفسها بغضّ النظر عن صورة الاختبار المقدمة له؛ فلذا تركّز بعض متطلبات المعادلة على درجات الطلبة، وكذلك فإنها تنطلق من منطلق أن توزيع درجات الطلبة الناتجة من الاختبار الأول هو توزيع الدرجات الناتجة من الاختبار الثاني نفسه؛ ولذا نجد أن بعض المتطلبات تركّز على توزيع درجات الأفراد، وتتحقّق هذه الشروط عند استعمال نظرية استجابة الفقرة إذا كانت البيانات تلائم النموذج المستعمل (Kolen & Brennan, 2014).

وقد اقترح علماء القياس والتقويم عددًا من المتطلبات لمعادلة درجات الاختبار، وعدم تحقّق هذه المتطلبات بالشكل المطلوب يؤدّي إلى الحكم بعدم تحقّق عملية المعادلة؛ ولذا فإن النتائج التي يُتَوَصَّل إليها عند عدم اكتمال تحقّق متطلبات المعادلة لا يمكن الوثوق بها أو الاعتماد عليها (الحربيّ وعليان، 2009)، وقد حدّد كلّ من علام (2005)، و(2014)Kolen and Brennan، و von Davier(2011)، و(1986) Croocker and Algina، و(2007) .Holland et al مجموعة من المتطلبات، ومنها:

1- التماثل Symmetry

يقصد به أن تكون تحويلات المعادلة متناظرة من صورة إلى أخرى وقابلة للانعكاس، بمعنى أنه يجب أن تكون دالة التحويل من درجات الصورة الأولى X إلى درجات الصورة الثانية Y هي عكس دالة التحويل من الصورة الثانية Y إلى الصورة الأولى X، وتعطي النتائج نفسها في الحالتين، فعلى سبيل المثال عند تحويل الدرجة χ على الصورة X إلى الدرجة γ على الصورة Y، فإنه يجب أن يكون تحويل الدرجة γ على الصورة Y إلى الدرجة χ على الصورة X هو نفسه. وتجدر الإشارة إلى أن معادلة الانحدار البسيط لا يمكن التعامل معها على أنها معادلة بين صور الاختبارات؛ وذلك لأن معادلة انحدار الصورة الأولى X على الصورة الثانية Y تختلف عن معادلة انحدار درجات الصورة الثانية Y على درجات الصورة الأولى X.

2- قياس المحتوى نفسه Same Specifications

ويقصد به أنه يجب أن تكون الصور المراد معادلتها تقيس المحتوى نفسه، وبُنِيَت وفقًا للمواصفات نفسها، ويُعَدّ هذه الشرط مهمًّا؛ حتى نتمكّن من استعمال الدرجات بين الصور المختلفة بالتبادل.

3- المساواة (التعادل) Equity

يقصد به أنه لا توجد فروق تعود إلى الفرد سواء أجاب عن الصورة الأولى أو الصورة الثانية من الاختبار، بمعنى أن يحصل المتقدّمون للامتحان على الدرجة المعادلة نفسها في الصورتين، فعلى سبيل المثال نفترض أن المتقدّمين للامتحان الذين لديهم درجة حقيقية معينة حصلوا في المتوسّط على درجة n على الصورة X، فسيحصل هؤلاء المتقدّمون أيضًا في المتوسّط على درجة معادلة n في الصورة Y؛ ولذا تكون الدرجات قابلة للتبادل بين الصورتين.

4-اللاتغير (عدم التباين في المجتمع) Group Invariance

بمعنى أن تكون دالة تحويل الدرجات أو معادلتها من الصورة الأولى إلى الصورة الثانية هي نفسها بغضّ النظر عن مجموعة المفحوصين الذين اختبروا، بمعنى استقلال دالة التحويل عن مجتمع المفحوصين، وكلما كانت صور الاختبار أكثر تشابهًا نتيجة بنائها وفقًا للمواصفات نفسها، كانت هذه الخاصية أكثر تحقّقًا.

5- الثبات Reliability

بمعنى يجب أن تتساوى صور الاختبارات المراد معادلتها في مستوى الثبات؛ لأنه لا يمكن القيام بإجراء عملية معادلة الدرجات على صور اختبارات تكون مختلفة في الثبات.

وتجدر الإشارة إلى أن هناك اختلافًا في وجهات نظر العلماء حول أهمية هذه المتطلبات، فمنهم من يرى أنها غير عملية أو صارمة بشكل كبير، وبصرف النظر عن الاختلافات في الآراء فإن هذه المتطلبات الخمسة لها قيمة كمعايير لتقييم ما إذا كان من الممكن معادلة صورتي اختبار بنجاح (Holland et al., 2007).

تصاميم جمع البيانات لإجراء المعادلة Equating data collection designs

إن عملية جمع البيانات تُعَدّ إحدى أهم ممارسات عملية المعادلة، ويشير Holland et al. (2007) إلى أن أحد الشواغل الرئيسة بين جميع أساليب معادلة الدرجات هو كيفية التحكّم في قدرات المفحوصين، وأوضح أنه إذا كان هناك اختباران مختلفان أُنْجِزا لمجموعتين مختلفتين من المفحوصين، فعند فحص توزيعات الدرجات فإن هناك عاملين يؤثّران في عملية تفسير النتائج: العامل الأول هو الصعوبة النسبية للاختبارين (وهو ما تهتم به عملية معادلة الاختبارات)، والعامل الآخر هو القدرة النسبية لمجموعات المفحوصين (وهو عامل غير مرغوب به، ويجب التخلّص منه أو التحكّم فيه)، وفي الممارسات العملية هناك طريقتان للفصل بين صعوبة الاختبار وقدرات المفحوصين: الطريقة الأولى

هي استعمال مجموعة مشتركة من المفحوصين وأحد أشكالها هو استعمال عينتين متكافئتين من المفحوصين، والطريقة الأخرى هي استعمال مقياس (اختبار) ثابت؛ لأن خضوع المفحوصين أنفسهم للاختبارين يتحكّم بشكل صريح في قدرة الممتحنين.

ولكلّ تصميم من تصاميم جمع البيانات مزايا وعيوب تجعله أكثر فائدة في مواقف وأقلّ فائدة في مواقف أخرى، فمثلًا يحتاج تصميم المجموعة الواحدة أحجامًا صغيرة، ويتطلّب تصميم المجموعات المتكافئة أحجامًا كبيرة؛ حتى تكون المعادلة بالدقة نفسها (von Davier, 2011)، وعند إجراء عملية المعادلة يجب أن تكون العينة التي تُطَبَّق الصور الاختبارية عليها ممثلة للمجتمع الأصلي، وتوجد عدة عوامل رئيسة يجب الاهتمام بها مثل ظروف التطبيق، وكذلك عامل التعب، وعامل التعلّم والتدريب؛ فلذا يجب الاهتمام بهذه العوامل عند اختيار التصميم المناسب (Kolen & Brennan, 2014)، وقد حدّد كلّ من Kolen and Brennan (2014)، وFelan (2002)، وvon Davier(2011)، وHolland et al. (2007) عدة تصاميم لإجراء المعادلة:

1- تصميم المجموعات العشوائية أو المتكافئة Random Groups Design

في هذا التصميم يُقَسَّم الممتحنون إلى مجموعتين: متماثلة ومتكافئة عشوائيًا، فعلى سبيل المثال تأخذ المجموعة الأولى الصورة الأولى وتأخذ المجموعة الثانية الصورة الثانية، وفي هذه الحالة يمكن حسبان الفرق بين الأداء على مستوى المجموعة على النموذجين مؤشّرًا مباشرًا للفرق بين الصعوبة في النموذجين، فمثلًا إذا كان متوسط درجات الممتحنين في الصورة الأولى 77، ومتوسط درجات الممتحنين في الصورة الثانية 72، فهذا يعني أن الصورة الأولى أسهل من الصورة الثانية.

ومن المميزات العملية لهذا التصميم أن كلّ مفحوص سيأخذ صورة واحدة من صور الاختبار؛ مما يؤدّي إلى تقليل وقت الاختبار مقارنة بالتصميم الذي يأخذ فيه المفحوص أكثر من صورة، وكذلك لا يتأثّر أداء المفحوص بعوامل مثل التعب أو التعلّم والتدريب، ويتطلّب هذا التصميم تطبيق جميع

الصور في الوقت نفسه مع المحافظة على ظروف التطبيق نفسها، والجدول (1) يوضّح تصميم المجموعات العشوائية (المتكافئة).

الجدول (1) تصميم المجموعات العشوائية (المتكافئة)

المجموعات	الصورة الأولى X	الصورة الثانية Y
1	✓	
2		✓

2 – تصميم المجموعة الواحدة Single Group Design

في هذا التصميم تعطى صورتا الاختبار للمجموعة نفسها من الممتحنين، بحيث تعطى إحدى الصورتين متبوعة بالصورة الأخرى، ويفضّل في اليوم نفسه؛ حتى لا تؤثر عوامل أخرى مثل عامل الخبرة في أداء الممتحنين في الاختبار، ويحصل كلّ ممتحن على تقديرين مختلفين للقدرة، لكلّ صورة تقدير مختلف، وفي هذه الطريقة لا تتداخل الفروق في مستويات الصعوبة بين الصورتين مع الفروق في مستويات قدرات المفحوصين، ويفترض أن عوامل مثل التعلّم والتدريب والخبرة ليس لها تأثير في أداء المفحوصين على الصورة الثانية من الاختبار، ويقترح وجود مدة زمنية؛ لتكون استراحة بين التطبيقين للحدّ من تأثير عامل التعب في أداء المفحوصين، ويجب الانتباه إلى أن الوقت الطويل بين التطبيقين سيؤدّي إلى تغيرات في قدرات المفحوصين، والجدول (2) يوضّح تصميم المجموعة الواحدة.

الجدول (2) تصميم المجموعة الواحدة

عدد المجموعات	الصورة الأولى X	الصورة الثانية Y
1	✓	✓

3- تصـــميم المجموعـة الواحـدة مع التوازن Single Group Design with Counterbalancing

هذا التصميم يساعد على التعامل مع تأثير ترتيب الصور في تصميم المجموعة الواحدة، ففي هذا التصميم يُقَسَّم الممتحنون إلى مجموعتين عشوائيًا، بحيث تأخذ المجموعة الأولى الصورة الأولى وتأخذ المجموعة الثانية الصورة الثانية في التطبيق الأول ثم يُعْكَس ترتيب الصور في التطبيق الثاني، والجدول (3) يوضّح تصميم المجموعة الواحدة المتوازنة.

الجدول (3) تصميم المجموعة الواحدة المتوازنة

المجموعات الفرعية	الصورة الأولى التطبيق الأول X_1	الصورة الأولى التطبيق الثاني X_2	الصورة الثانية التطبيق الأول Y_1	الصورة الثانية التطبيق الثاني Y_2
1	✓			✓
2		✓	✓	

4- تصـــميم الجـذع المشـــترك للمجموعـات غير المتكـافئـة Common-Item Nonequivalent Groups Design

يستعمل هذا التصميم عندما لا يمكن تطبيق أكثر من صورة للاختبار في الوقت نفسه، ففي هذا التصميم يُقَسَّم الممتحنون إلى مجموعتين، بحيث تأخذ المجموعة الأولى الصورة الأولى وتأخذ المجموعة الثانية الصورة الثانية، ثم يُطَبَّق اختبار مشترك على المجموعتين، ويمكن أن يكون اختبار الجذع المشترك داخليًا، بحيث تكون فقراته مضمنة ضمن صورتي الاختبار المراد معادلتهما، وقد يكون اختبار الجذع المشترك خارجيًا، وذلك عندما يُتَعامَل معه كاختبار مستقلّ يُقَدَّم للممتحنين خارج الوقت المحدّد للصورتين المراد معادلتهما.

5- تصميم الجذع المشترك Anchor-test Groups Design

وفيه تُطَبَّق صورتا الامتحان على مجموعتين من المفحوصين بالتصميم السابق نفسه، ولكن الفرق هو أن مجموعات المفحوصين متكافئة، والجدول (4) يوضّح ذلك.

الجدول (4) تصميم الجذع المشترك

المجموعات	الصورة الأولى X	الجذع المشترك A	الصورة الثانية Y
1	✓	✓	
2		✓	✓

وتجدر الإشارة إلى أن لكل تصميم شروطًا يجب تحقّقها، وفيما يلي توضيح للتصميم المستعمل في هذه الدراسة وهو تصميم المجموعات العشوائية (المتكافئة) Random Groups Design، حيث يُعَدّ خيارًا مناسبًا للعينات ذات الأحجام الكبيرة، وكذلك فإنه يساعد على تجنّب مشكلة تأثير ترتيب إعطاء الصور الامتحانية التي قد تنشأ في تصميم المجموعة الواحدة عندما يأخذ كلّ مفحوص كلا الصورتين (von Davier, 2011)، ويشير Dorans et al. (2007) إلى أنه في تصميم المجموعات العشوائية يمكن حسبان الفرق في الأداء بين المجموعات بمثابة مؤشر مباشر للفرق في الصعوبة بين الصورتين.

طرائق معادلة درجات الاختبارات

ذكر كلّ من Kolen and Brennan (2014)، و Felan (2002)، و Holland et al. (2007)، و von Davier (2011)، و De Ayala (2009)، و Hambleton and Swaminathan (1985) عدة طرائق لمعادلة درجات الاختبارات منها ما يستند إلى النظرية الكلاسيكية للقياس، ومنها ما يستند إلى نظرية الاستجابة للفقرة.

1- طرائق معادلة الدرجات باستعمال النظرية الكلاسيكية Classical Test Theory CTT

- **طريقة معادلة الوسط الحسابي Mean Equating**

في هذا الطريقة يمكن إجراء تحويل رياضي لجعل الوسط الحسابي لدرجات الصورة الأولى مساويًا للوسط الحسابي لدرجات الصورة الثانية، وتقوم هذه الطريقة على افتراض تساوي مستويات الصعوبة بين الصورتين Kolen and Brennan (2014)، ويمكن التعبير عنها رياضيًا بالصورة الآتية:

$$X_1 - \mu_{X_1} = X_2 - \mu_{X_2}$$

حيث X_1 هي درجة الطالب في الصورة الأولى، وX_2 هي درجة الطالب في الصورة الثانية.

μ_{X_1} يمثل الوسط الحسابي لدرجات الصورة الأولى، و μ_{X_2} يمثل الوسط الحسابي لدرجات الصورة الثانية.

- **طريقة المعادلة الخطية Linear Equating**

هذه الطريقة تفترض أن الاختلاف بين صور الاختبار يكون في المتوسطات الحسابية والانحرافات المعيارية، حيث يُحْسَب المتوسط الحسابي والانحراف المعياري لدرجات الطلبة على الصورتين؛ ولذا تكون الدرجات المتكافئة على صورتي الاختبار هي الدرجات التي تكون لها الدرجة المعيارية Z Score نفسها (الحربيّ وعليان، 2009). ويعبر عنها رياضيًا:

$$\frac{X - \mu_X}{S_X} = \frac{Y - \mu_Y}{S_Y}$$

$$Y = \left(\frac{S_Y}{S_X}\right) \mathrm{X} + \left(\mu_Y - \frac{S_Y}{S_X}\mu_X\right)$$

$$\mathrm{Y} = \mathrm{A} * \mathrm{X} + \mathrm{B}$$

حيث $B = \mu_Y - \frac{S_Y}{S_X}\mu_X$ ، $A = \frac{S_Y}{S_X}$

وX هي درجة الطالب في الصورة الأولى Y هي درجة الطالب في الصورة الثانية

و μ_X متوسط درجات الصورة الأولى، و S_X الانحراف المعياري لدرجات الصورة الأولى،

و μ_Y متوسط درجات الصورة الثانية، و S_Y الانحراف المعياري لدرجات الصورة الثانية.

ومن الطريقة الخطية يمكن اشتقاق طرائق أخرى منها طريقة معادلة المنحنى التربيعي، وطريقة توكر، وطريقة ليفين، وطريقة المعادلة الدائرية (Kolen & Brennan, 2014).

- **المعادلة بواسطة الرتب المئينية Equipercentile Equating**

في هذه الطريقة تكون الدرجات متعادلة في صورتي الاختبار إذا كان لديها الرتب المئينية نفسها، ويُنْجَز ذلك بجعل التوزيع الإحصائي لدرجات إحدى الصورتين بعد التحويل مماثلًا للتوزيع الإحصائي لدرجات الصورة الثانية، بحيث يكون له المتوسط الحسابي والانحراف المعياري نفسه؛ ولذا تكون الرتب المئينية المتساوية لدرجات المفحوصين في صورتي الاختبار تدلّ على تساوي مستوى الأداء في كلتا الصورتين، وهذا هو أساس معادلة درجات الاختبار بهذه الطريقة (علام، 2005).

2- طرائق المعادلة وفقا لنظرية الاستجابة للفقرة Item Response Theory (IRT)

نظرًا لإخفاق طرائق النظرية الكلاسيكية للقياس لمعادلة الدرجات الخام في تحقيق بعض شروط المعادلة مثل العدالة أو اللاتباين في المجتمع أو التماثل؛ فإنها أصبحت غير مرغوبة أحيانًا؛ لذا فإن معادلة درجات الامتحانات باستعمال نظرية الاستجابة للفقرة تحلّ الكثير من المشكلات التي عجزت عن حلّها النظرية الكلاسيكية في القياس، بشرط أن تكون الدرجات المراد معادلتها متطابقة مع النموذج المستعمل (المدانات والكيلاني، 2008).

وحدّد كلّ من Kolen and Brennan (2014)، و Hambleton and Swaminathan (1985) عددًا من الخطوات التي يجب اتباعها عند إجراء معادلة درجات الاختبارات باستعمال نظرية الاستجابة للفقرة، وهي:

1. تحديد تصميم جمع البيانات المناسب، وذلك حسب طبيعة الاختبارات المراد معادلة درجاتها وخصائص المفحوصين.

2. تحديد نموذج نظرية الاستجابة للفقرة المناسب، والذي يتطابق مع البيانات، وذلك عن طريق مقاييس جودة المطابقة.

3. وضع تقديرات المعلمات على تدريج مشترك، بحيث يربط بين القدرة المراد قياسها ومعلمة صعوبة الفقرة.

4. تسجيل درجات المفحوصين على صورتي الاختبار حسب التدريج المناسب، فمثلًا يمكن تسجيل درجات المفحوصين كدرجات حقيقية، أو كدرجات قدرة، أو كدرجات ملاحظة.

وتوجد ثلاث طرائق لمعادلة درجات الاختبارات باستعمال نظرية الاستجابة للفقرة كما أشار إليها كلّ من: Kolen and Brennan (2014)، و Hambleton and Swaminathan (1985):

- **معادلة الدرجات باستعمال الدرجات الحقيقية True Score equating**

تتضمن هذه الطريقة تحديد درجات حقيقية متكافئة على كل من صورتي الاختبار عن طريق استعمال دالة تربط بين تقديرات القدرة θ والعلامات الحقيقية للقدرة، وتصاغ رياضيًا على الصورة:

$$\zeta = \sum_{i=1}^{n} P_i(\theta)$$

حيث تشير المعادلة إلى أنه يمكن تحديد الدرجة الحقيقية للمفحوص ذي القدرة θ على الصورة X بحساب احتمالية إجابته إجابة صحيحة عن الفقرة $P_i(\theta)$ لجميع فقرات الاختبار، ثم تُجْمَع هذه الاحتمالات لجميع الفقرات للحصول على الدرجة الحقيقية للاختبار (المدانات والكيلاني، 2008).

فإذا افترضنا أن صورتي الاختبار تقيسان السمة نفسها، ووُضِعَت معالم الفقرات على تدريج مشترك، فإن الدرجتين الحقيقيتين مرتبطتان بدرجات القدرة بواسطة منحنى خصائص الاختبار (الحربيّ وعليان، 2009).

- **معادلة الدرجات باستعمال درجات القدرة Ability Score equating**

في هذه الطريقة يفترض أن كلا الامتحانين يقيسان السمة نفسها، فإذا قدرنا معالم القدرة ومعالم الفقرات في الوقت نفسه لصورتي الاختبار، ستكون معالم القدرة ومعالم الفقرات على التدريج نفسه، فستكون هناك علاقة خطية بين التدريجين؛ ولذا يمكن إجراء المعادلة بين صورتي الاختبار (الحربيّ وعليان، 2009).

- **معادلة الدرجات باستعمال الدرجات الملاحظة Observed Score equating**

تقوم هذه الطريقة على التنبؤ بالتوزيع النظريّ للدرجات الملاحظة (الدرجات الخامّ) للاختبار عن طريق بناء التوزيع التكراريّ للدرجات الخامّ لمفحوص قدرته θ ، وتُعَدّ هذه الطريقة أقل تعقيدًا من طريقة الدرجات الحقيقية في عملية المعادلة؛ لأنها لا تنظر بشكل مباشر للمتغيرات الكامنة، ولا تتطلّب أن تكون صور الامتحانات المراد معادلتها قد بنيت بجدول المواصفات نفسه(Kolen & Brennan, 2014) .

طرائق حساب معاملات تحويل المقاييس (معاملات الربط):

عند استعمال نماذج نظرية الاستجابة للفقرة لتحليل صور اختبارية متعددة، تظهر النتائج مختلفة باختلاف هذه الصور مالم يُرْبَط بينها؛ فلذا تستعمل طرائق الربط لتحويل المعلمات إلى مقياس مشترك يسمح بمقارنة الأفراد عبر الصور الاختبارية المختلفة، وتفسير النتائج في ضوء معيار موحد، وذلك بتحويل التدريج المستقلّ لمفردات الصورتين إلى تدريج واحد مشترك، وذلك بعد التحقّق من تحقيق بيانات الصورتين لافتراضات نظرية الاستجابة للفقرة (الوليلي، 2005).

وهناك طريقتان لعمل مقياس مشترك، فالمقياس المشترك يمكن عمله عن طريق التقدير المتزامن لمعالم الفقرات لبيانات صورتي الامتحان، وتُسَمّى بالمعايرة المتزامنة، أو يمكن عمله عن طريق تقدير معالم الفقرة لصورتي الاختبار بتشغيلين منفصلين، ثم القيام بعملية الربط باستعمال طرائق الربط

(المتوسط/المتوسط، المتوسط/ الانحراف المعياري، ستوكنج- لورد، هيبارا) لوضع المعالم على مقياس مشترك (Song, 2009).

1- طريقة المتوسط - المتوسط (Mean/ Mean): تُعَدّ هذه الطريقة من الطرائق المبسطة، وذلك بحساب التحويل الخطي باستعمال متوسط تقديرات معلمة الصعوبة b ومعلمة التمييز a في صورتي الاختبار، حيث يمكن تقدير الثابتين (معاملات الربط) A، B باستعمال المعادلات الآتية:

$$A = \frac{\bar{a}_Y}{\bar{a}_X}$$

$$B = \bar{b}_Y - A\bar{b}_X$$

حيث إن: $\bar{a}_Y$، $\bar{a}_X$ تمثل متوسطات معلمات التمييز في الصورتين X, Y و $\bar{b}_Y$، $\bar{b}_X$ تمثل متوسطات معلمات الصعوبة في الصورتين X, Y (Xu, 2009).

2- طريقة المتوسط - الانحراف المعياري (Mean/ Sigma): في هذه الطريقة تُسْتَعْمَل المتوسطات والانحرافات المعيارية لتقديرات معلمة الصعوبة b، وتكون هذه الطريقة مناسبة عند اختلاف معلمة التمييز بين الفقرات (Baker & Kim, 2004)، ويكون ذلك عن طريق التحويل:

$$A = \frac{\sigma_{b_X}}{\sigma_{b_Y}}$$

$$B = A\bar{b}_Y - \bar{b}_X$$

حيث $\bar{b}_Y$، $\bar{b}_X$ تمثل متوسطات معلمات الصعوبة في الصورتين X, Y و σ_{b_Y}، σ_{b_X} تمثل الانحرافات المعيارية لمعلمات الصعوبة في الصورتين X, Y (Zhonghua, 2010).

3- **طريقة ستوكنج لورد** Stocking-Lord: تُعَدّ هذه الطريقة من طرائق تحويل المنحنيات المميزة للفقرة، حيث اقْتُرِحَت هذه الطرائق لتجنّب بعض المشكلات في طريقتي المتوسط/ المتوسط و المتوسط/ الانحراف المعياري، ففي هذه الطريقة يمكن إيجاد معاملات التحويل الخطيّ لمعالم الفقرات لإحدى مجموعتي الممتحنين، بحيث تتطابق هذه المعالم مع دالة خصائص الاختبار للفقرات المشتركة بين

الصورتين، وفي هذه الطريقة يمكن إيجاد معاملات التحويل (الميل والجزء المقطوع من المحور الصادي) لصعوبة الفقرات وتمييزها في مجموعة واحدة عن طريق مطابقة منحنيات خصائص الاختبار (Von Davier, 2011).

4- **طريقة هيبارا** Haebara: تُعبِّر هذه الطريقة عن الاختلافات بين منحنيات خصائص الفقرة كمجموع مربعات الاختلافات بين دوال خصائص الفقرة لكلّ فقرة من الفقرات المشتركة للممتحنين من قدرة معينة(Von Davier, 2011).

ثانيًا: الدراسات السابقة:

يُعَدّ مجال معادلة درجات الاختبارات باستعمال نظرية الاستجابة للفقرة محطّ اهتمام الكثير من الباحثين في مجال القياس والتقويم؛ وذلك بسبب المزايا التي تقدمها هذه النظرية وقدرتها على معالجة جوانب القصور التي كانت تعاني منها طرائق معادلة الدرجات باستعمال النظرية الكلاسيكية.

حيث أُنْجِزَت مجموعة من الدراسات منها ما يهتمّ بمعادلة صور اختبارية مختلفة، ومنها ما تطرّق إلى المقارنة بين طرائق المعادلة، وقد اهتمّ بعضها بتأثير بعض خصائص العينة في دقّة المعادلة، وفيما يلي عرض لأبرز هذه الدراسات:

فقد هدفت دراسة الشريفين (2009) إلى معادلة درجات سبعة نماذج مختلفة لاختبار الكفاءة في اللغة الإنجليزية الذي عقدته جامعة اليرموك في العام الدراسي 2006 / 2007م، وقد كانت العينة جميع الطلبة الذين تقدّموا للاختبار والبالغ عددهم 1199 طالبًا وطالبة. قام الباحث باستعمال نموذج راش لمعادلة درجات النماذج السبعة من الاختبار كنموذج من نماذج نظرية الاستجابة للفقرة، واستعمل طريقة المعادلة الخطية والمعادلة المئينية كطرائق لنظرية الاستجابة للفقرة، وأظهرت هذه الدراسة أن النماذج السبعة لم تكن متكافئة في أوساطها الحسابية وتبايناتها، وقد بيّنت نتائج هذه الدراسة أن درجة القدرة المقابلة للدرجة الخام في النموذج الأول والتي تُعَدّ درجة النجاح تقابلها درجات قدرة متفاوتة تراوحت بين −0.54 و 0.01 لوجيت، وأوصت هذه الدراسة باستعمال الدرجات المتعادلة في اتخاذ قرار النجاح والرسوب، وليس على أساس الدرجات الخام.

وهدفت دراسة حمادنة (2012) إلى الكشف عن أثر تجانس المحتوى وطول الجذع المشترك في دقّة معادلة درجات اختبار ثنائي التدريج ومتعدّد المستوى لمادة الرياضيات باستعمال صورتين من اختبارات الجذع المشترك، واحْتُسِبَت دقّة المعادلة من خلال الخطأ المعياري للمعادلة المئينية، وتكوّنت عينة الدراسة من 776 طالبًا وطالبة من الصفين: التاسع والعاشر بالأردن، حيث اخْتِيروا بطريقة

عشوائية، وقد أظهرت نتائج هذه الدراسة أن دقّة المعادلة العمودية كانت أعلى ما يمكن عندما كان اختبار الجذع المشترك متجانسًا وعدد فقراته أكثر من 20 فقرة.

وأنجز حمادنة (2013) دراسة هدفت إلى الكشف عن أثر وجود أداء تفاضلي في الفقرات المرساوية (الجذع المشترك) في دقّة المعادلة العمودية لاختبار أوتيس لينون للقدرة العقلية بالمستوى المتوسط والمتقدم للصورة K والمعدلتين على البيئة الأردنية، حيث تكوّنت عينة الدراسة من 1001 طالب وطالبة، منهم 500 في المستوى القدرة المتوسط و501 في مستوى القدرة المتقدم، ولعملية جمع بيانات الدراسة فقد وظّف الباحث تصميم الفقرات المشتركة والمجموعات غير المتكافئة، وتضمّنت الصورتان 25 فقرة مشتركة، وللكشف عن الأداء التفاضلي للفقرات تبعًا للجنس بشكل منفصل لكلّ صورة فقد وظّف الباحث طريقة مانتل هانزل، ثم قام بإجراء المعادلة العمودية بين درجات الصورتين وفقًا للنموذج أحادي المعلمة من نماذج نظرية الاستجابة للفقرة، حيث قام بعملية المعادلة مرتين: المرة الأولى مع الاحتفاظ بالفقرات المشتركة ذات الأداء التفاضلي والمرة الأخرى مع حذف الفقرات المشتركة ذات الأداء التفاضلي، وللمقارنة بين عمليتي المعادلة استعمل الباحث معامل الصدق التقاطعي معيارًا لتحديد مدى استقرار النتائج لدرجات القدرة المعدلة ومؤشرًا على فعالية المعادلة. أظهرت نتائج هذه الدراسة إلى أن حذف الفقرات المشتركة ذات الأداء التفاضلي أدى إلى زيادة دقّة المعادلة العمودية بالاعتماد على قيم معاملات الصدق التقاطعي.

وقام He et al. (2013) بدراسة هدفت إلى فحص تأثير إزالة القيم المتطرفة في معادلة الدرجات الحقيقية باستعمال طريقة الانحدار الخطي، حيث وظّف الباحثون طريقة المعادلة باستعمال الدرجات الحقيقية وتصميم الفقرات المشتركة والمجموعات غير المتكافئة، واستعمل الباحثون نموذجين للاختبار أحدهم مكوّن من 120 فقرة من فقرات الاختيار من متعدّد مع 40 فقرة مشتركة، والاختبار الثاني مكوّن من 30 فقرة من فقرات الاختيار من متعدّد مع 10 فقرات مشتركة، حيث وُلِّدَت استجابات للفقرات تتلاءم مع النموذج ثلاثي

المعلمة بحجم عينتين 1000 و 5000 على التوالي، وقد أمكن القيام بعملية المعادلة باستعمال الدرجات الحقيقية لنظرية الاستجابة للفقرة، وأظهرت نتائج هذه الدراسة أن طريقة الكشف عن الفقرات المشتركة المتطرفة باستعمال الانحدار الخطي كان أداؤها أفضل من الطريقة التقليدية في معظم الظروف.

وهدفت دراسة المحرزيّ (2014) إلى معادلة أربعة نماذج من اختبار القدرة اللفظية واختبار القدرة الكمية من اختبارات القدرة العامة الذي يُعِدّه المركز الوطني للقياس والتقويم في وزارة التعليم العالي بالمملكة العربية السعودية باستعمال ثلاث طرائق من طرائق النظرية الكلاسيكية للمعادلة (المعادلة الخطية و المعادلة المئينية والمعادلة المئينية الممهدة بعديًا)، وبتوظيف تصميم المجموعات المتكافئة والمقارنة بين هذه الطرائق، وشملت الدراسة طلاب الصفين: الأول الثانوي والثاني الثانوي بجميع مناطق المملكة العربية السعودية للعام الأكاديمي 1430 / 1431هـ، وبلغ حجم عينة الدراسة التي طُبِّقَت عليها نماذج الاختبار 1081 و 1082 و 1083 و 2034 طالبًا وطالبة، حيث أظهرت نتائج الدراسة عدم تكافؤ النماذج الأربعة للاختبارين، والحاجة إلى معادلة هذه النماذج، وأظهرت النتائج أيضًا أن الدرجات المعادلة الناتجة من الطرائق الثلاث مختلفة، وأن طريقة المعادلة المئينية الممهدة حقّقت أقل مقدار للأخطاء المعيارية عبر جميع درجات الاختبار.

وهدفت دراسة عبدالوهاب (2015) إلى الكشف عن أثر اختلاف معامل ثبات صورتي الاختبار المراد معادلتهما في دقة المعادلة عند استعمال طرائق المعادلة القائمة على نظرية الاستجابة للفقرة، ولتحقيق أهداف الدراسة وظّف الباحث صورتين لاختبار مقرر سيكولوجية التعلّم في الفصل الأول من العام الجامعي 2014 / 2015م، وبلغ حجم عينة الدراسة 951 طالبًا وطالبة من طلبة الفرقة الثانية بكلية التربية بجامعة المنيا في العام الجامعي 2014/ 2015م، وأظهرت نتائج هذه الدراسة إلى أن دقّة المعادلة تقلّ بزيادة الفرق بين معاملي الثبات في حالة استعمال طريقتي هايبارا وستوكنج لورد والعكس.

وهدفت دراسة المحرزيّ (2015) إلى مقارنة الدرجات المكافئة لنماذج اختبار القدرات العامة في تصميم المفردات المشتركة بمجموعات غير متكافئة باستعمال خمس طرائق من طرائق المعادلة القائمة على النظرية الكلاسيكية، وهي المعادلات الخطية (براون هولاند، لفين – توكر)، والمعادلات غير الخطية (المئينية والمئينية الممهدة بعديا)، وقد قام الباحث بمعادلة نموذجين جديدين مع نموذج مرجعي في اختباري القدرة اللفظية والقدرة الكمية عن طريق عينة تكوّنت من طلاب الصفين: الثاني الثانوي والثالث ثانوي من مناطق المملكة العربية السعودية المختلفة في نهاية العام الأكاديمي 1430/ 1431 هـ، وقد بلغ حجم العينة 6948 طالبًا وطالبة، وأظهرت نتائج هذه الدراسة أن الدرجة المتكافئة الناتجة من طرائق المعادلة الخطية تتشابه مع الدرجات الناتجة من طرائق المعادلة غير الخطية، وأظهرت النتائج أيضًا أن الطريقة المئينية حصلت على أصغر متوسط للأخطاء المعيارية للمعادلة.

وهدفت دراسة Wiberg (2016) باستعمال المنهج التجريبي إلى فحص ثلاث طرائق خطية بديلة لطريقة معادلة الدرجات الملاحظة لنظرية الاستجابة للفقرة، ومقارنة هذه الطرائق البديلة مع الطريقة التقليدية لمعادلة الدرجات الملاحظة لنظرية الاستجابة للفقرة، وذلك بتوظيف تصميم المجموعات المتكافئة، وباستعمال دراسة محاكاة وبيانات حقيقية من اختبار التقييم الدراسي السويدي(SweSAT) ، وهو اختبار جامعي يتكوّن من 160 فقرة من فقرات الاختيار من متعدّد موزّعة على ثمانية اختبارات فرعية مقسّمة إلى قسمين: لفظيّ وكميّ، ولدراسة المحاكاة أُعِدَّت صورتين للاختبار، وتحتوي كلّ صورة على 40 فقرة، وشملت الدراسة 82000 مفحوص لكلّ صورة من صور الاختبار، حيث قُسِّمَ المفحوصون على مستويات القدرة $\theta=-2$ إلى $\theta=2$، فكلّ مستوى قدرة ضمّ 2000 مفحوص. وأظهرت نتائج الدراسة أن الطرائق البديلة حققت أداءً جيدًا مقارنة بالطرائق التقليدية والطريقة الخطية مع وجود استثناءات قليلة.

وأنجز الرواجفة والطراونة (2016) دراسة لإجراء المعادلة العمودية لمعادلة اختبار متعدّد المستويات للصفوف: العاشر، والأول الثانوي العلمي، والثاني الثانوي العلمي باستعمال نظرية الاستجابة

للفقرة، وباستعمال طريقة المعادلة بدرجات القدرة وفق نموذج راش، فقد بُنِيَت ثلاثة اختبارات تحصيلية، وتكّون كلّ اختبار من 25 فقرة مع وجود فقرات مشتركة بين كلّ اختبارين، وتكوّنت عينة الدراسة من 800 طالب وطالبة موزّعين على 35 مدرسة حكومية من مدارس قصبة الطفيلة بالأردن للعام الدراسي 2014 / 2015م، وأظهرت نتائج الدراسة فعالية المعادلة العمودية في معادلة امتحانات المستويات الثلاثة عن طريق قيم معاملات الصدق التقاربي.

وهدفت دراسة المحروق (2017) إلى معرفة دقّة معادلة درجات الاختبارات متعددة الاستجابة في ضوء اختلاف حجم العينة ومستوى الصعوبة بتوظيف نظرية الاستجابة للفقرة وتصميم المجموعات العشوائية، وبلغ حجم العينة 1800 موزّعة على ثلاثة مستويات، وذلك بتوليد بيانات تجريبية باستعمال برمجية (Wingen2)، وأظهرت نتائج الدراسة ان حجم العينات الكبيرة يقلّل من الخطأ المعياري للمعادلة؛ مما يؤدّي إلى زيادة دقّة المعادلة، وأظهرت النتائج أيضًا أن النماذج المتشابهة في صعوبتها تميل فيها قيم الخطأ المعياري للانخفاض والعكس.

وهدفت دراسة الحبسيّ (2018) إلى معادلة درجات امتحان دبلوم التعليم العامّ لمادة الفيزياء في سلطنة عُمان بتوظيف طريقة معادلة الدرجات الحقيقية المستندة إلى نظرية الاستجابة للفقرة باستعمال تصميم المجموعات المتكافئة. وشملت الدراسة جميع طلبة الدبلوم العام والذين خضعوا لاختبار مادة الفيزياء خلال العامين الدراسيين: 2014 / 2015م والعام 2015 / 2016م، والبالغ عددهم 18966 و17175 طالبًا وطالبة على التوالي، وتكوّنت أداة الدراسة من صورتين لامتحان مادة الفيزياء للصف الثاني عشر، فقد تألّفت الصورة الأولى من 40 فقرة، وتألّفت الصورة الثانية من 36 فقرة، وأظهرت نتائج الدراسة تقارب المتوسطات الحسابية للدرجات الحقيقية للصورة المرجعية والدرجات المعادلة للصورة الجديدة.

وهدفت دراسة الزعابية وحسن (2018) إلى معادلة كتيبات اختبار الدراسة الدولية لقياس مهارات القراءة PIRLS للعام الدراسي 2011م في سلطنة عُمان باستعمال تصميم المفردات المشتركة

والمجموعات غير المتكافئة، وذلك بتوظيف طريقتين من طرائق نظرية الاستجابة للفقرة هما المعادلة باستعمال الدرجات الحقيقية والمعادلة باستعمال الدرجات الملاحظة. تكوّن الاختبار من 146 فقرة اختبارية منها 74 مفردة ثنائية الاستجابة و72 مفردة متعدّدة الاستجابة موزّعة على 13 كتيبًا وُزِّعَت على 13 مجموعة من طلبة الصف الرابع الأساسي بسلطنة عُمان بعينة بلغ حجمها الإجمالي 10394 طالبًا وطالبة، وأظهرت نتائج الدراسة وجود اختلاف بين القيم المعادلة بطريقتي المعادلة بالدرجات الحقيقية والمعادلة بالدرجات الملاحظة لصالح المعادلة بالدرجات الملاحظة في جميع الكتيبات ما عدا الكتيب 10، فقد كان الاختلاف فيه لصالح المعادلة بالدرجات الحقيقية، وأمّا في الكتيبين: 4 و5، فلا يوجد اختلاف بين القيم المعادلة بالطريقتين.

وقام العنزيّ (2018) بإجراء دراسة هدفت إلى التعرّف على أثر طول الجذع المشترك في دقّة معادلة صورتي الاختبار في مادة الرياضيات وفق تصميم الجذع المشترك بتوظيف طريقة معادلة الدرجات الحقيقية وطريقة معادلة الدرجات المشاهدة. وقد قام الباحث ببناء صورتين متكافئتين لاختبار تحصيليّ في مادة الرياضيات يتكوّن كلّ منهما من 60 فقرة من فقرات الاختيار من متعدّد، واختبار مشترك لاختبار تحصيليّ في مادة الرياضيات، وجميع فقراته من نوع الاختيار من متعدّد، وقد أخذ متغيّر طول الجذع المشترك ثلاثة مستويات هي: 30، و20، و10 فقرة اختبارية، وقد طُبِّقَ على عينة عشوائية من طلبة الصف الثالث متوسط للعام الدراسي 2013/ 2014م بمنطقة تبوك بالمملكة العربية السعودية.

أظهرت نتائج هذه الدراسة أنه كلما قلّ طول الجذع المشترك قلّ المتوسّط الحسابي للقيم المعادلة باستعمال الدرجات الحقيقية، وأنه لا توجد فروق بين متوسطات قيم الدرجات المتعادلة باستعمال طريقة المعادلة باستعمال الدرجات الملاحظة، وأشارت نتائج الدراسة أيضًا إلى دقّة المعادلة عند استعمال جذع مشترك بطول 20 فقرة؛ فقد كانت أكثر دقّة من استعمال جذع مشترك بطول 10 فقرات في الطريقتين.

وأجرت العبيدي (2018) دراسة هدفت إلى معادلة اختبارات فليب كارتر الأربعة المتقدمة في الذكاء باستعمال الدرجة الحقيقية وفق النظرية الكلاسيكية، ولتحقيق أهداف الدراسة وظّفت الباحثة المنهج الوصفي التحليلي، فقد طبّقت اختبارين من سلسلة اختبارات فيليب كارتر وكين راسل للقدرة العقلية على عينة بحجم 560 طالبًا وطالبة من طلبة المرحلة الإعدادية بمديرية محافظة بغداد في العراق، وأظهرت نتائج الدراسة أن الاختبارين متعادلان بدرجة كبيرة وفق الطرائق الثلاث التي اتبعتها الباحثة في عملية المعادلة، وهي طريقة المتوسط الحسابي وطريقة المعادلة الخطية وطريقة المعادلة المئينية.

وهدفت دراسة الشريفين وأبي العناز (2019) إلى تحديد أثر حجوم العينات الصغيرة في دقّة المعادلة والمقارنة بين طريقة كيرنل (النواة) وطريقة قوس الدائرة في معادلة درجات اختبار على عينات صغيرة، ولتحقيق أهداف الدراسة أعدّ الباحثان صورتين لاختبار تحصيلي على وحدة التكامل من كتاب الرياضيات للصف الثاني الثانوي بعدد فقرات 25 فقرة لكلّ منهما من فقرات الاختيار من متعدّد بأربعة بدائل لكل فقرة باستعمال تصميم المجموعات المتكافئة بعينة ذات حجوم مختلفة: (15، و25، و50، و100) طالبة من طالبات الصف الثاني الثانوي بمحافظة عجلون بالأردن للعام الدراسي 2014/ 2015م، وأظهرت نتائج الدراسة أن طريقة قوس الدائرة كانت أكثر دقّة من طريقة كيرنل (النواة)، وأظهرت النتائج أيضًا أن الأخطاء المعيارية تقلّ بزيادة حجم العينة لكلا الطريقتين.

وهدفت دراسة الزعابية وآخرين (2020) إلى المقارنة بين طريقة معادلة الدرجات الحقيقية وطريقة معادلة الدرجات المشاهدة في إجراء معادلة كتيبات الاختبارات المكوّنة من فقرات ثنائية الاستجابة ومتعدّدة الاستجابة باستعمال تصميم المجموعة الواحدة، وتصميم المجموعات غير المتكافئة ذات الفقرات المشتركة، ومن أجل تحقيق غرض الدراسة فقد وظّف الباحثون اختبار قياس مهارة التقدّم في القراءة PIRLS لعام 2011م، والذي يتكوّن من 146 فقرة اختبارية منها 74 مفردة ثنائية الاستجابة و72 مفردة متعدّدة الاستجابة موزّعة على 13 كتيّبًا وُزِّعَت على 13 مجموعة من طلبة الصف الرابع الأساسي

بسلطنة عُمان بعينة بلغ حجمها الإجمالي 10394 طالبًا وطالبة، وقد استعملت طريقتين من طرائق المعادلة باستعمال نظرية الاستجابة للفقرة: طريقة معادلة الدرجات الحقيقية، وطريقة معادلة الدرجات المشاهدة، وأوضحت نتائج الدراسة تقارب الإحصائيات الوصفية لكتيبات الاختبار بعد المعادلة بالطريقتين، وأوضحت النتائج أيضًا وجود فروق بين متوسطات القيم المعادلة بطريقتي الدرجات الحقيقية والدرجات الملاحظة لصالح الدرجات الملاحظة لجميع الكتيبات ما عدا الكتيب 10 الذي كانت الفروق بين المتوسطات فيه لصالح المعادلة بالدرجات الحقيقية، وأمّا الكتيبان: 4 و5، فقد أظهرا عدم وجود فروق.

وهدفت دراسة جاد الله وآخرين (2021) إلى التعرّف على أثر اختلاف طريقة المعادلة في القيم الناتجة من عملية المعادلة وكذلك أثرها في دقّة المعادلة حيث تم استخدام تصميم المجموعات غير المتكافئة ذات الجذع المشترك، والذي تطلّب إعداد صورتين لامتحان تحصيلي في مادة العلوم للصف الثاني الإعدادي وبين الصورتين بعض الفقرات المشتركة، وشملت عينة البحث 1032 من طلبة الصف الثاني الإعدادي الأزهري بالقاهرة في العام الدراسي 2020/ 2021م، وقد قُسِّموا إلى مجموعتين غير متكافئتين، ولكلّ مجموعة إحدى الصورتين، وأظهرت الدراسة وجود فروق في القيم الناتجة من عملية المعادلة لصالح طريقة المتوسط/ الانحراف المعياري؛ فهي أعطت أعلى قيم، وأظهرت النتائج أيضًا وجود فروق في دقّة المعادلة لصالح طريقة المتوسط/ الانحراف المعياري؛ فقد كانت أكثر دقة.

وأجرى العنزيّ (2021) دراسة هدفت إلى تحديد أثر تصميم المعادلة المستعمل في دقّة المعادلة، فقد وظّف تصميم الفقرات المشتركة، وتصميم المجموعات المتكافئة وتصميم الأفراد المشتركين باستعمال طرائق نظرية الاستجابة للفقرة، ولتحقيق أهداف الدراسة استعمل الباحث امتحان ثنائي البعد في مادة الحساب، وطُبِّقَ على عينة مكوّنة من 1240 من طلبة المرحلة المتوسط بدولة الكويت في العام الدراسي 2019/ 2020م، وأظهرت نتائج الدراسة وجود فروق في دقّة المعادلة تبعًا للتصميم المستعمل لصالح تصميم الفقرات المشتركة؛ فقد كان أكثر كفاءة ودقّة في المعادلة.

وهدفت دراسة Leoncio et al. (2023) إلى المقارنة بين ثلاث طرائق من طرائق لمعادلة الدرجات هي طريقة الدرجة الملاحظة لنظرية الاستجابة للفقرة، وطريقة النواه للنظرية الكلاسيكية، وطريقة النواة لنظرية الاستجابة للفقرة، وقد تطلّب إجراء هذه الدراسة توليد بيانات تسمح بمحاكات درجات الاختبار، وأظهرت نتائج الدراسة أن طريقتي الاستجابة للفقرة تقدّم نتائج أفضل من طريقة النظرية الكلاسيكية، ولكنها أسرع في عملية المعادلة من طرائق نظرية الاستجابة؛ ولذا فهي مناسبة للتطبيقات المتكررة.

وهدفت دراسة القرني (2023) إلى تحديد أثر طريقة معالجة القيم المفقودة وتقدير قدرة الفرد على دقّة معادلة درجات الاختبارات باستعمال نظرية الاستجابة للفقرة، وقد وظّف الباحث اختبار القدرة العقلية لـ أوتيس ولينون، وذلك بتطبيقه على 3276 من طلبة المرحلة المتوسطة بمدارس مدينة جدة بالمملكة العربية السعودية، وقام الباحث باستعمال ثلاث طرائق لمعالجة القيم المفقودة هي: خوارزمية تعظيم التوقعات، وحساب قيم تعويضية متعددة، وحساب التعويض باستعمال دالة الاستجابة للفقرة، وكذلك استعمل أربع طرائق لتقدير قدرات الأفراد هي: الأرجحية العظمى، والتوقّع البعدي، والأرجحية العظمى الموزونة، والطريقة البايزية، وأظهرت نتائج الدراسة أن أفضل طريقة لتعويض القيم المفقودة هي طريقة دالة الاستجابة للفقرة، وذلك عند استعمال طريقة التوقّع البعدي لتقدير قدرات الأفراد، وقدّمت طريقة تقدير القدرات بالأرجحية العظمى الموزونة أفضل دقّة لتقدير القدرات عند وجود بيانات مفقودة، وأظهرت النتائج أيضًا أن أفضل طريقة لتعويض البيانات المفقودة عند تقدير قدرات الأفراد باستعمال الأرجحية العظمى هي طريقة التعويض بدالة الاستجابة للفقرة.

وقام Fellinghauer et al. (2023) بدراسة محاكاة هدفت إلى معرفة أثر الاختلاف في البناء المقاس وصعوبة الاختبار في دقّة معادلة الدرجات عند استعمال تصميم المجموعات المشتركة، فقد قامت الدراسة بمحاكات درجات اختبارات مع تغيير عدد الفقرات وأحجام العينة، وأظهرت النتائج أن

الاختلاف في البناء المقاس يؤدّي إلى خفض دقّة المعادلة، وأن حجم العينة الأكبر لم يحسّن دقّة المعادلة، بينما المقاييس الأطول حسّنت دقّة المعادلة.

وأجرى Wallmark et al. (2023) دراسة محاكاة للبيانات الحقيقية هدفت إلى المقارنة بين طريقة معادلة النواه وطريقة المعادلة بتوظيف الدرجة الملاحظة للاختبارات المختلطة وباستعمال نظرية الاستجابة للفقرة بتصميم المجموعات المتكافئة وتصميم الجذع المشترك للمجموعات غير المتكافئة. أظهرت نتائج هذه الدراسة أن الفرق بين طريقة معادلة النواة وطريقة المعادلة بالدرجات الحقيقية كان قليلًا سواء في الدرجات المعادلة أو في الأخطاء المعيارية للمعادلة.

وهدفت دراسة الحجوري وعبد الوهاب (2024) إلى معرفة أثر نموذج نظرية الاستجابة للفقرة في دقّة معادلة الاختبارات، وركزت بالتحديد على النموذج ثنائي البارامتر والنموذج ثلاثي البارامتر، وذلك باستعمال طريقة المتوسط / المتوسط وبتصميم الجذع المشترك، حيث قام بتوليد بيانات محاكاة لأربع صور اختبارية لكلّ نموذج صورتين، وأظهرت نتائج الدراسة أن النموذج الثنائي أكثر دقّة في معادلة الدرجات من النموذج الثلاثي عند استعمال طريقة المتوسط/ المتوسط.

وقام Sun and Kim (2024) بدراسة محاكاة هدفت إلى تحديد تأثير الاختلاف في الصعوبة في دقّة معادلة الدرجات عند استعمال تصميم المجموعات العشوائية و تصميم الجذع المشترك للمجموعات غير المتكافئة، وقام الباحث بتقييم أداء ست طرائق من طرائق معادلة الدرجات، وكشفت نتائج هذه الدراسة أنه في تصميم المجموعات العشوائية كانت طريقة المعادلة بالمتوسط الحسابي أكثر دقّة عندما تكون الصورتان متساويتين في الصعوبة أو يكون الفرق في الصعوبة صغيرًا، بينما تكون طريقة المعادلة بالنسبة المئوية أكثر دقّة عندما يكون الفرق في الصعوبة بين الصورتين كبيرًا، وعند استعمال تصميم الجذع المشترك للمجموعات غير المتكافئة فطريقة توكر الخطية أكثر دقّة عندما يكون الفرق في الصعوبة صغيرًا أو متوسطًا، بينما مع الفرق الكبير في الصعوبة يفضل استعمال طريقة النسب المئوية المتساوية.

وأجرى Wallin and Wiberg (2025) دراسة محاكاة هدفت إلى المقارنة بين أثر طريقة تنعيم البيانات في دقّة معادلة الدرجات باستعمال طرائق نظرية الاستجابة للفقرة والطرائق القائمة على النظرية الكلاسيكية، وذلك بالمقارنة بين النماذج الشائعة لعملية التنعيم ونموذج التنعيم ثنائي المتغير باستعمال تصميم الجذع المشترك والمجموعات غير المتكافئة، وأظهرت النتائج أن نموذج التنعيم المسبق يؤثر في الدرجات المتعادلة في الطرائق القائمة على النظرية الكلاسيكية، ولكن بالنسبة للطرائق القائمة على نظرية الاستجابة للفقرة فقد تحقّقت دقّة عالية عن طريق نماذج التنعيم المختلفة.

مناقشة الدراسات السابقة:

من استعراض الدراسات السابقة نجد أن الأهداف الرئيسة لهذه الدراسات قد تنوّعت؛ فلذا يمكن تقسيمها حسب هدفها العام إلى ثلاثة أقسام رئيسة هي:

- دراسات هدفت إلى معادلة صور اختبارية باستعمال نظرية الاستجابة للفقرة، ومنها: الشريفين (2009)، والمحرزي (2014)، والرواجفة والطراونة (2016)، والحبسيّ (2018)، والزعابية وحسن (2018).
- دراسات هدفت إلى المقارنة بين طرائق معادلة درجات الاختبار باستعمال نظرية الاستجابة للفقرة، ومنها: المحرزي (2015)، و(2016) Wiberg، والزعابية وآخرون (2020)، وجاد الله وآخرون (2021)، و(2023) Leoncio.
- دراسات هدفت إلى المقارنة بين تصاميم جمع البيانات ونموذج نظرية الاستجابة المستعمل، ومنها العنزيّ (2021)، والحجوري وعبد الوهاب (2024)، و(2025) Wallin and Wiberg.
- دراسات هدفت لدراسة العوامل المؤثرة في دقة المعادلة مثل تصميم جمع البيانات، وحجم العينة، وطول الاختبار، وصعوبة الفقرات، وطريقة المعادلة، ومنها: حمادنة (2012)، والمحروق (2017)، والعنزيّ (2018)، والشريفين وأبو العناز (2019).

حيث أوضحت هذه الدراسات الاهتمام المتزايد بموضوع معادلة درجات الاختبارات، وبسبب تنوّع هذه الدراسات واختلافها في طرائق المعادلة، وكذلك اختلاف الصور الاختبارية وكذلك اختلافها في تصاميم جمع البيانات ومنهجياتها، فقد اعتمدت بعض الدراسات على بيانات واقعية، وبعضها الآخر كانت دراسات محاكاة اعتمدت على بيانات توليدية، فإنه يصعب التوصل إلى تعميمات دقيقة حول نتائج تلك الدراسات.

وتتفق الدراسة الحالية مع بعض الدراسات في كونها هدفت إلى معادلة صور امتحانية مختلفة، حيث تتشابه مع دراسات كل من: الشريفين (2009)، والمحرزي (2014)، والرواجفة والطراونة (2016)، والزعابية وحسن (2018)، والعبيدي (2018).

وتتشابه الدراسة الحالية مع دراسة الحبسيّ (2018) في أنها هدفت إلى معادلة درجات امتحان دبلوم التعليم العام، لكن تختلف معها في أن هذه الدراسة تخصّ اختبار مادة الرياضيات المتقدمة، وأنها تختصّ بامتحانات مناهج سلاسل كامبريدج؛ فإن منهجية الامتحانات في هذه المناهج تختلف عن منهجية الامتحانات في المناهج السابقة.

وتُعَدّ الدراسة الحالية هي أول دراسة لمعادلة امتحان دبلوم التعليم العام لمناهج سلاسل كامبريدج بشكل عامّ ومادة الرياضيات المتقدمة بشكل خاصّ في سلطنة عُمان.

واستفادت الدراسة الحالية من الدراسات السابقة في التعرّف على طرائق عملية معادلة درجات الاختبارات وخطواتها باستعمال نظرية الاستجابة للفقرة، وكذلك التعرّف على البرامج الحاسوبية المستعملة في عملية معادلة درجات الاختبارات.

الفصل الثالث

منهجية الدراسة وإجراءاتها

أولًا: منهجية الدراسة

- منهج الدراسة
- مجتمع الدراسة وعينتها
- أداة الدراسة
- تصميم الدراسة

ثانيًا: إجراءات الدراسة والمعالجات الإحصائية

- إجراءات الدراسة
- المعالجات الإحصائية
  - المعالجات الإحصائية للإجابة عن السؤال الأول
  - المعالجات الإحصائية للإجابة عن السؤال الثاني
  - المعالجات الإحصائية للإجابة عن السؤال الثالث
  - المعالجات الإحصائية للإجابة عن السؤال الرابع
  - المعالجات الإحصائية للإجابة عن السؤال الخامس
  - المعالجات الإحصائية للإجابة عن السؤال السادس

الفصل الثالث

منهجية الدراسة وإجراءاتها

يتناول هذا الفصل توضيح منهجية الدراسة، بالإضافة إلى وصف لمجتمع الدراسة وعينتها، وكذلك توضيح لمكوّنات أداة الدراسة، ويتناول أيضًا خطوات معادلة درجات صورتي امتحان الدبلوم العام لمادة الرياضيات المتقدمة باستعمال نظرية الاستجابة للفقرة، وسيُتَطَرَّق أيضًا في هذا الفصل إلى الإجراءات والطرائق الإحصائية المتبعة في الإجابة عن أسئلة الدراسة.

منهج الدراسة

تستعمل الدراسة الحالية المنهج الوصفي؛ وذلك لملائمته لطبيعة الدراسة وأهدافها، حيث يعتمد هذا المنهج على وصف الظاهرة، وتحديد خصائصها، وجمع البيانات حولها، ثم تحليل هذه البيانات وتفسيرها لاستخلاص النتائج، وستُحَلَّل نتائج امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة للعامين الدراسيين: 2023/ 2024م و2024/ 2025م ومعادلتها باستعمال نظرية الاستجابة للفقرة .

مجتمع الدراسة وعينتها

يتكوَّن مجتمع الدراسة الحالية وعينتها من جميع طلبة دبلوم التعليم العامّ في جميع محافظات سلطنة عُمان النظاميين في المدارس الحكومية والخاصّة الذين خضعوا لامتحان مادة الرياضيات المتقدمة للفصل الدراسي الأول للعامين الدراسيين: 2023/ 2024م و2024/ 2025م، والبالغ عددهم (24619) طالبًا وطالبة و(28659) طالبًا وطالبة على التوالي (وزارة التربية والتعليم، 2025). ويوضح الجدول (5) توزيع أعداد الطلبة على الصورتين حسب المحافظة التعليمية والنوع الاجتماعي.

الجدول (5) أعداد المتقدمين لامتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة بسلطنة عُمان للعامين الدراسيين: 2023/ 2024م و2024/ 2025م وفق المحافظة والنوع الاجتماعي لصورتي الامتحان

المحافظة التعليمية	الصورة الأولى 2024/2023م			الصورة الثانية 2024/ 2025م		
	ذكور	إناث	المجموع	ذكور	إناث	المجموع
مسقط	2516	2785	5301	3022	3477	6499
شمال الباطنة	1934	2281	4215	2227	2691	4918
الداخلية	2027	1987	4014	2257	2205	4462
جنوب الباطنة	1387	1641	3028	1626	1820	3446
الظاهرة	710	677	1387	759	853	1612
شمال الشرقية	877	1051	1928	1013	1079	2092
جنوب الشرقية	946	1091	2037	1066	1318	2384
ظفار	932	789	1721	1102	1016	2118
البريمي	238	286	524	361	346	707
الوسطى	110	144	254	99	120	219
مسندم	94	116	210	89	113	202
المجموع	11771	12848	24619	13621	15038	28659

أداة الدراسة

- تتكوّن أداة الدراسة من صورتين لامتحانين لمادة الرياضيات المتقدمة لدبلوم التعليم العام للفصل الدراسي الأول للعامين الدراسيين: 2023/ 2024م و2024/ 2025م، حيث تتألّف الصورة الأولى من (27) فقرة، وتتألف الصورة الثانية من (29) فقرة، وجميعها إلزامية الإجابة، وكلّ فقرة تقيس مخرجًا تعليميًا واحدًا فقط لا يتكرّر، وقد أُعِدَّت الصور الامتحانية وفق وثيقة تقويم تعلّم الطلبة في مادة الرياضيات للصفين: الحادي عشر والثاني عشر من فريق مختصّ في المادة التعليمية على مستوى وزارة التربية والتعليم، وطُبِّقَ في نهاية الفصل الدراسي، وحُدِّدَ زمن

الإجابة عن امتحان الدبلوم العام لمادة الرياضيات المتقدمة بثلاث ساعات، واشتملت كلّ صورة امتحانية على ثلاثة أنواع من الفقرات:

– مفردات الاختيار من متعدّد: تضمّ كلّ صورة اختبارية 12 فقرة من فقرات الاختيار من متعدّد لكلّ منها درجة واحدة، وتوضع الفقرة في أيّ مكان في الورقة الامتحانية، بمعنى لا يوجد قسم منفصل لفقرات الاختيار من متعدّد.

– مفردات قصيرة الإجابة: هي مفردات إجابة نصيّة حرّة خُصِّص لكلّ منها من 2–3 درجات.

– مفردات طويلة الإجابة: وتمثّل فقرات يجب فيها على الطالب القيام بعدّة خطوات للوصول إلى الحلّ النهائي، وخُصِّص لكلّ منها 5 درجات.

حيث روعي عند إعداد الامتحان النهائي الأوزان النسبية للمخرجات التعليمية للمنهج والمستويات التعليمية لها، ويرتكز التقويم في مناهج كامبريدج العالمية على هدفين:

1– هدف التقويم الأول: (AO1: Assessment Objective) المعرفة والفهم: ويركز على المفاهيم والمهارات والإجراءات التي يتعلّمها الطلبة، ويشتمل على قدرات [التذكر، التعريف، إجابات قصيرة، الحساب، الاستخراج] (وزارة التربية والتعليم، 2024ب).

2– هدف التقويم الثاني: (AO2: Assessment Objective2) التطبيق والاستدلال: ويركز على قدرة الطلبة على تطبيق المفاهيم والمهارات والإجراءات في حلّ المسائل أو المشكلات الرياضية، ويشتمل على قدرات [الاختيار، حلّ المسائل، التنفيذ، النمذجة، التركيب والدمج، التبرير والبرهان، حلّ المشكلات] (وزارة التربية والتعليم، 2024ب).

وتُصَنَّف الفقرات الامتحانية سواء على الهدف الأول أو الثاني للتقويم حسب صعوبتها إلى منخفضة (L) وتضم 40% من فقرات الامتحان بواقع 28 درجة من إجمالي درجات الامتحان النهائي،

ومتوسطة (M) وتضم 40% من فقرات الامتحان بواقع 28 درجة من إجمالي درجات الامتحان النهائي، ومرتفعة (H) وتضم 20% من فقرات الامتحان بواقع 14 درجة من إجمال درجات الامتحان النهائي، ويهدف هذا التصنيف إلى إتاحة الفرصة لتفريد التعلم بين الطلبة ومراعات الفروق الفردية (وزارة التربية والتعليم، 2024ب).

ويقوم فريق من المحكّمين بخلاف الفريق المختصّ بإعداد الامتحان بمراجعة فقرات الامتحان وإبداء ملاحظاتهم حول دقّة صياغة الفقرات وارتباطها بالمحتوى التعليمي للمادة الدراسية، وتمثيلها التمثيل الجيد لمستويات التعلّم حسب جدول المواصفات؛ ولذا فإن هذه الصور الامتحانية تتوافر بها أدلة صدق المحتوى حسب جداول المواصفات التي أعدّها المختصّون في وزارة التربية والتعليم.

ويوضح الجدول (6) توزيع درجات الامتحان النهائي لمادة الرياضيات المتقدمة حسب هدفي التقويم ومستويات صعوبة الفقرة، وكذلك توزيع الدرجات على الوحدات الدراسية (وزارة التربية والتعليم، 2024ب).

الجدول (6) توزيع درجات الامتحان النهائي لمادة الرياضيات المتقدمة حسب هدفي التقويم ومستوى الصعوبة والوحدة الدراسية

الوحدة الدراسية		القياس الدائري		حساب المثلثات		مقدمة في النهايات والاتصال		التفاضل
توزيع الدرجات		13 درجة		22 درجة		16 درجة		19 درجة
الدرجة الكلية		70 درجة						
أهداف التقويم		هدف التقويم الأول AO1 50%				هدف التقويم الثاني AO2 50%		
توزيع الدرجات	الدرجة الكلية للهدف	35 درجة				35 درجة		
	مستوى الصعوبة	منخفض الصعوبة	متوسط الصعوبة	مرتفع الصعوبة	منخفض الصعوبة	متوسط الصعوبة	مرتفع الصعوبة
	درجة كل مستوى	40% 14 درجة	40% 14 درجة	20% 7 درجات	40% 14 درجة	40% 14 درجة	20% 7 درجات

ويوضح الجدولان: (7، و8) توزيع فقرات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام على هدفي التقويم ومستوى صعوبة الفقرة (مركز القياس والتقويم، اتصال شخصي، مايو11، 2025).

الجدول (7) توزيع فقرات امتحان الصورة الأولى 2024/2023م

الفقرة	الوحدة	هدف التقويم	مستوى الصعوبة	نوع الفقرة	الدرجة
1	القياس الدائري	AO1	منخفضة	ثنائية	1
2	حساب المثلثات	AO1	منخفضة	ثنائية	1
3	مقدمة في النهايات والاتصال	AO1	منخفضة	متدرجة	4
4	مقدمة في النهايات والاتصال	AO1	منخفضة	ثنائية	1
5	التفاضل	AO1	منخفضة	ثنائية	1
6	التفاضل	AO1	منخفضة	متدرجة	5
7	التفاضل	AO1	منخفضة	ثنائية	1
8	القياس الدائري	AO1	متوسطة	متدرجة	3
9	حساب المثلثات	AO1	متوسطة	ثنائية	1
10	حساب المثلثات	AO1	متوسطة	متدرجة	4
11	مقدمة في النهايات والاتصال	AO1	متوسطة	ثنائية	1
12	التفاضل	AO1	متوسطة	متدرجة	5
13	حساب المثلثات	AO1	مرتفعة	متدرجة	3
14	مقدمة في النهايات والاتصال	AO1	مرتفعة	متدرجة	3
15	التفاضل	AO1	مرتفعة	ثنائية	1
16	القياس الدائري	AO2	منخفضة	متدرجة	4
17	حساب المثلثات	AO2	منخفضة	ثنائية	1
18	مقدمة في النهايات والاتصال	AO2	منخفضة	ثنائية	1
19	التفاضل	AO2	منخفضة	متدرجة	4
20	التفاضل	AO2	منخفضة	متدرجة	4
21	حساب المثلثات	AO2	متوسطة	متدرجة	5
22	مقدمة في النهايات والاتصال	AO2	متوسطة	متدرجة	5
23	التفاضل	AO2	متوسطة	ثنائية	1
24	التفاضل	AO2	متوسطة	متدرجة	3
25	حساب المثلثات	AO2	مرتفعة	متدرجة	3
26	مقدمة في النهايات والاتصال	AO2	مرتفعة	متدرجة	3
27	التفاضل	AO2	مرتفعة	ثنائية	1
		المجموع			70

الجدول (8) توزيع فقرات امتحان الصورة الثانية 2024/ 2025م

الفقرة	الوحدة	هدف التقويم	مستوى الصعوبة	نوع الفقرة	الدرجة
1	القياس الدائري	AO1	منخفضة	ثنائية	1
2	حساب المثلثات	AO1	منخفضة	ثنائية	1
3	حساب المثلثات	AO1	منخفضة	متدرجة	3
4	مقدمة في النهايات والاتصال	AO1	منخفضة	ثنائية	1
5	مقدمة في النهايات والاتصال	AO1	منخفضة	متدرجة	4
6	التفاضل	AO1	منخفضة	متدرجة	4
7	القياس الدائري	AO1	متوسطة	متدرجة	2
8	حساب المثلثات	AO1	متوسطة	ثنائية	1
9	حساب المثلثات	AO1	متوسطة	متدرجة	3
10	مقدمة في النهايات والاتصال	AO1	متوسطة	ثنائية	1
11	مقدمة في النهايات والاتصال	AO1	متوسطة	متدرجة	2
12	التفاضل	AO1	متوسطة	ثنائية	1
13	التفاضل	AO1	متوسطة	متدرجة	4
14	التفاضل	AO1	مرتفعة	ثنائية	1
15	القياس الدائري	AO1	مرتفعة	متدرجة	3
16	حساب المثلثات	AO1	مرتفعة	متدرجة	3
17	القياس الدائري	AO2	منخفضة	متدرجة	3
18	حساب المثلثات	AO2	منخفضة	ثنائية	1
19	حساب المثلثات	AO2	منخفضة	متدرجة	4
20	مقدمة في النهايات والاتصال	AO2	منخفضة	ثنائية	1
21	التفاضل	AO2	منخفضة	متدرجة	5
22	القياس الدائري	AO2	متوسطة	متدرجة	3
23	القياس الدائري	AO2	متوسطة	ثنائية	1
24	حساب المثلثات	AO2	متوسطة	متدرجة	5
25	مقدمة في النهايات والاتصال	AO2	متوسطة	متدرجة	4
26	التفاضل	AO2	متوسطة	ثنائية	1
27	حساب المثلثات	AO2	مرتفعة	ثنائية	1
28	مقدمة في النهايات والاتصال	AO2	مرتفعة	متدرجة	3
29	التفاضل	AO2	مرتفعة	متدرجة	3
المجموع					70

تصميم الدراسة

اعتمدت الدراسة الحالية على تصميم المجموعات العشوائية أو المتكافئة Random Groups Design كما هو موضح في الجدول (9)، فقد استعملت درجات صورتين من امتحان مادة الرياضيات المتقدمة لعامين دراسيين مختلفين للمحتوى نفسه (محتوى الفصل الدراسي الأول)، وقد طُبِّقَت الصورة الأولى X على مجموعة من المفحوصين هم طلبة الدبلوم العام للعام الدراسي 2023/ 2024م، وطُبِّقَت الصورة الثانية Y على مجموعة أخرى هم طلبة الدبلوم العام للعام الدراسي 2024/ 2025م.

الجدول (9) تصميم المعادلة المستعمل في الدراسة

المجموعة	الصورة X	الصورة Y
طلبة الدبلوم العام لعام 2023/ 2024م	✓	
طلبة الدبلوم العام لعام 2024/ 2025م		✓

إجراءات الدراسة

للإجابة عن أسئلة الدراسة قام الباحث بمجموعة من الخطوات الإجرائية، فقد اتبع الإجراءات والخطوات المتمثلة في الآتي:

1. الاطلاع على الأدب النظري في مجال معادلة الصور الامتحانية واختيار التصميم المناسب لجمع البيانات لمعادلة صورتي الامتحان، واختار تصميم المجموعات المتكافئة أو العشوائية، حيث قدمت صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام لمجموعتين عشوائيتين من الطلبة بحسبان أن توزيع الطلبة على العامين الدراسيين كان عشوائيًا.
2. الحصول على خطاب تسهيل مهمة باحث من جامعة الشرقية إلى وزارة التربية والتعليم (ملحق 1).
3. تقديم طلب الحصول على البيانات اللازمة للدراسة إلى وزارة التربية والتعليم (ملحق 2).

4. الحصول على موافقة وزارة التربية والتعليم على السماح بالحصول على البيانات والمعلومات المتعلقة بالدراسة (مركز القياس والتقويم، اتصال شخصي، أبريل 21، 2025).
5. الحصول على بيانات (درجات) صورتي امتحان مادة الرياضيات المتقدمة لدبلوم التعليم العام للعامين الدراسيين: 2023/ 2024م والذي يمثل الصورة الأولى، و2024/ 2025م والذي يمثل الصورة الثانية.
6. تهيئة البيانات؛ لتتلاءم مع البرمجيات المستعملة.
7. تحديد دلالات صدق البناء والثبات لصورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة لدبلوم التعليم العام للعامين الدراسيين: 2023/ 2024م، و2024/ 2025م.
8. التحقق من مدى مطابقة درجات صورتي الامتحان النهائي للفصل الدراسي الأول لمادة الرياضيات المتقدمة لدبلوم التعليم العام لافتراضات نظرية الاستجابة للفقرة.
9. التحقق من مدى مطابقة فقرات صورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة لنموذج نظرية الاستجابة للفقرة: النموذج ثلاثي المعالم **(3PLM)** لفقرات الاختيار من متعدد، ونموذج التقدير الجزئي المعمم **(G−PCM)** للفقرات المقالية، حيث إن نموذج التقدير الجزئي يناسب وصف استجابات اختبارات التحصيل مثل الرياضيات، فتكون الإجابة الصحيحة جزئيا ممكنة (Embretson & Reise, 2000)، وكذلك فإنه مناسب عندما يأخذ المصححون في الحسبان مجموعة من مكونات الاستجابة المختلفة، ويجعلون لكلّ منها ما هو صحيح أو غير صحيح، وتكون درجة الفقرة هي عدد الخطوات التي أُكْمِلَت بنجاح (Reckase, 2009)، ويشير Bezirhan and von Davier (2023) إلى أن نموذج التقدير الجزئي المعمم (G−PCM) يستعمل في تقدير معالم فقرات اختبارات TIMSS و PIRLS.

10. معايرة فقرات صورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة باستعمال (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية) وتقدير معالم الفقرة (الصعوبة والتمييز والتخمين).
11. التحقق من مدى مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة وتقدير قدرات الأفراد.
12. معادلة درجات صورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة لدبلوم التعليم العام.

المعالجات الإحصائية

للإجابة عن أسئلة الدراسة وتحقيق أهدافها فقد استعملت البرمجيات: (برنامج R حزمة mirt، SPSS, Excel) لتحليل البيانات واستخراج النتائج، فقد أُنْجِزَت المعالجات الإحصائية الآتية:

- للإجابة عن السؤال الأول "**ما دلالات صدق البناء والثبات لصورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة لدبلوم التعليم العام للعامين الدراسيين 2023/ 2024م، و2024/ 2025م؟**"

1. حُسِبَت الإحصاءات الوصفية (المتوسط والانحراف المعياري للدرجات الخام لصورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام، وحُسِبَت قيم التفلطح والالتواء للتحقق من التوزيع الطبيعي لبيانات الصورتين.
2. استعمال الاختبار الإحصائي (اختبار ت) للعينات المستقلة لتحديد طبيعة الفروق بين المتوسطات الحسابية للصورتين.
3. التحقق من صدق البناء والثبات لصورتي الامتحان، وقد تُحقَّق من الصدق بالتحليل العاملي الاستكشافي، والتحليل العاملي المتوازي، ففي البداية تم التحقق من صلاحية بيانات الصورتين

للتحليل العاملي من خلال الإحصائي Kaiser–Meyer–Olkin (KMO)، ثُمَ تم إجراء التحليل العاملي الاستكشافي والمتوازي لبيانات الصورتين.

وكذلك تم التحقق من الثبات لصورتي الامتحان من خلال معامل ثبات ألفا كرونباخ، ومعامل ماكدونالد أوميجا.

- للإجابة عن السؤال الثاني " **ما مدى مطابقة درجات صورتي الامتحان النهائي للفصل الدراسي الأول لمادة الرياضيات المتقدمة لدبلوم التعليم العام لافتراضات نظرية الاستجابة للفقرة؟"**

1. تُحُقِّق من أحادية البُعد لبيانات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال التحليل العاملي الاستكشافي، والتحليل العاملي المتوازي.
2. تُحُقِّق من الاستقلال الموضعي لبيانات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال مصفوفة البواقي المعيارية.

- للإجابة عن السؤال الثالث "**ما مدى مطابقة فقرات صورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة لنموذج نظرية الاستجابة للفقرة (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية)؟"**

1. تُحُقِّق من مدى مطابقة بيانات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام لنماذج نظرية الاستجابة للفقرة عن طريق تحليل بيانات الصورتين معًا باستعمال حزمة (mirt)، فقد تم التحقق من عملية المطابقة عن طريق اختبار M2، وكذلك عن طريق تحقّق مؤشرات جودة المطابقة.
2. كذلك تُحُقِّق من مطابقة بيانات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال الإحصائيات outfit و infit.

- للإجابة عن السؤال الرابع "**ما تقدير معالم فقرات صورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة باستعمال (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية؟**"

1. قُدِّرت معالم الفقرات ثنائية الاستجابة لصورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال (النموذج ثلاثي المعالم (3PLM).
2. قُدِّرت معالم الفقرات المقالية لصورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال نموذج التقدير الجزئي المعمم (G–PCM).

- للإجابة عن السؤال الخامس "**ما مدى مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة وتقدير قدرات الأفراد؟**"

1. التحقّق من مدى مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة باستعمال مؤشر المطابقة infit، فقد حُسِبَت نسبة الأفراد المطابقين لنماذج نظرية الاستجابة للفقرة لمجموعتي طلبة دبلوم التعليم العام.
2. حُسِبَت الإحصاءات الوصفية لتوزيع قدرات مجموعتي طلبة دبلوم التعليم العام.

- للإجابة عن السؤال السادس "**ما قيم الدرجات المتعادلة لصورتي الامتحان النهائي لمادة الرياضيات المتقدمة لدبلوم التعليم العام؟**"

1. تهيئة درجات صورتي الامتحان النهائي لمادة الرياضيات المتقدمة لدبلوم التعليم العام لإجراء المعادلة بين درجات الصورتين، فقد دُمِجَت 4 فقرات ثنائيًا من فقرات امتحان الصورة الثانية؛ وذلك بهدف جعل عدد الفقرات في الصورتين متساويًا، وهو (27) فقرة لكل صورة؛ فلذا دُمِجَت الفقرة (7) مع الفقرة (9)، ودُمِجَت الفقرة (3) مع الفقرة (11)، وجاءت عملية الدمج مع مراعاة أن تكون الفقرتان على هدف التقويم نفسه، وأيضًا من مستوى الصعوبة نفسه، وأيضًا مع مراعاة أن لا تتجاوز الدرجة الكلية للفقرة بعد عملية الدمج أعلى درجة للفقرات الأصلية للامتحان.

2. إجراء عملية المعادلة بين صورتي الامتحان النهائي لمادة الرياضيات المتقدمة لدبلوم التعليم العام باستعمال برنامج R، فقد اعتمدت الصورة الأولى 2023/ 2024م لتكون هي الصورة المرجعية.
3. إجراء عملية معادلة درجات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال الدرجات الخام.
4. إجراء عملية معادلة درجات صورتي الامتحان النهائي لمادة الرياضيات المتقدمة لدبلوم التعليم العام باستعمال درجات القدرة.
5. إجراء عملية معادلة الدرجات الخام للصورة الأولى لامتحان الرياضيات المتقدمة لدبلوم التعليم العام بما يقابلها من درجات قدرة، ودرجات حقيقية على امتحان الصورة الثانية.

الفصل الرابع

نتائج الدراسة ومناقشتها

أولًا. عرض نتائج الدراسة ومناقشتها:

- نتائج السؤال الأول ومناقشتها
- نتائج السؤال الثاني ومناقشتها
- نتائج السؤال الثالث ومناقشتها
- نتائج السؤال الرابع ومناقشتها
- نتائج السؤال الخامس ومناقشتها
- نتائج السؤال السادس ومناقشتها
- ملخّص النتائج

ثانيًا: المقترحات والتوصيات

الفصل الرابع

نتائج الدراسة ومناقشتها

مقدمة

يتضمّن هذا الفصل تقديم عرض للنتائج التي توصلت إليها الدراسة ومناقشتها في ضوء الأدبيات والدراسات السابقة، وذلك بالإجابة عن أسئلة الدراسة، حيث تعتمد دقة النتائج على دقة البرامج المستعملة، وسَتُعرَض كذلك المقترحات والتوصيات وفقًا لنتائج الدراسة الحالية.

النتائج المتعلقة بالسؤال الأول ومناقشتها

للإجابة عن السؤال الأول الذي ينصّ على "ما دلالات صدق البناء والثبات لصورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة لدبلوم التعليم العام للعامين الدراسيين 2023/ 2024م، و2024/ 2025م؟"، فقد حُسِبَت الإحصاءات الوصفية لصورتي الامتحان، ثم تُحُقِّق من صدق البناء والثبات للصورتين، فبالنسبة للإحصاءات الوصفية يوضح الجدول (10) الإحصاءات الوصفية لصورتي امتحان الرياضيات المتقدمة بالدرجات الخامّ.

الجدول (10) الإحصاءات الوصفية لصورتي امتحان الرياضيات المتقدمة وقيم التفلطح والالتواء

صورة الامتحان	حجم العينة	المتوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح	قيمة ت	مستوى الدلالة	حجم التأثير
الصورة الأولى 2023/ 2024م	24619	51.79	16.68	−0.88	−0.29	14.90	0.001	0.13
الصورة الثانية 2024/ 2025م	28659	49.53	18.47	−0.66	−0.81			

من الجدول (10) يتضح أن المتوسط الحسابي لدرجات الصورة الأولى كان أعلى من متوسط درجات الصورة الثانية، حيث بلغ متوسطها 51.79، وبانحراف معياري 16.68، وأن قيم التفلطح والالتواء للصورتين

كانت ضمن النطاق الطبيعي لمستويي التفلطح والالتواء (−3 إلى +3) (George & Mallery, 2010) ، مما يشير إلى أن درجات صورتي امتحان الرياضيات المتقدمة تتوزّع توزيعًا طبيعيًا.

وللكشف عن طبيعة الفروق في المتوسطات الحسابية لصورتي الامتحان اسْتُعْمِل اختبار (ت) للعينات المستقلة، حيث يلاحظ من الجدول (10) وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية للدرجات الخامّ في صورتي امتحان الرياضيات المتقدمة لصالح الصورة الأولى، حيث كان لها أعلى متوسط وبحجم تأثير قليل (كامل، 2022)؛ حيث إن الاختلاف في المتوسطات بين النماذج يعود إلى الاختلاف في صعوبتها أو بسبب اختلاف قدرات الأفراد (المحرزي، 2015)، وهذا يؤكد ضرورة إجراء معادلة لدرجات الصورتين، وتتفق هذه النتيجة مع دراسات كلّ من: (جاد الله وآخرين، 2021؛ الحبسيّ، 2018؛ الزعابية وحسن، 2018؛ الشريفين، 2009، العنزي؛ 2021، المحرزي، 2014؛ Wiberg, 2016).

وبعد التحقق من الإحصاءات الوصفية لصورتي الامتحان تُحقِّق من الصدق البنائي والثبات لصورتي امتحان الرياضيات المتقدمة.

ويُعَدّ التحليل العاملي الاستكشافي Exploratory Factor Analysis (EFA) أحد إجراءات التحقق من صدق البناء، حيث تُحقِّق من الافتراضات الإحصائية للتحليل العاملي بواسطة الاختبار الإحصائي Kaiser–Meyer–Olkin (KMO)، ويوضح الجدول (11) التحقق من محكات صلاحية بيانات صورتي الامتحان النهائي في مادة الرياضيات المتقدمة لدبلوم التعليم العام للتحليل العاملي.

الجدول (11) التحقق من محكات صلاحية بيانات صورتي الامتحان للتحليل العاملي

الصورة	كايزر – ماير KMO	مربع كاي Chi–Square	درجات الحرية	مستوى الدلالة	محدد مصفوفة معاملات الارتباط
الأولى 2023 / 2024م	0.98	281882.8	351	0.001	0.00001
الثانية 2024/ 2025م	0.99	411861.6	406	0.001	0.00001

من الجدول (11) نلاحظ تحقق افتراضات التحليل العاملي لصورتي الامتحان، وهذه الافتراضات هي:

- عدم وجود مشـــكلة التداخل الخطي المتعدّد (Multi–Collinearity) بين المتغيرات: فقد كانت قيمة محدد المصفوفة (Determinant) للصورتين على التوالي 0.00001، 0.00001، حيث إن هذه القيم صـغيرة، وفي الوقت نفسـه لا تسـاوي صـفرًا، وهذا يشـير إلى عدم وجود مشـكلة التداخل الخطي بين المتغيرات (Beavers et al.,2013).
- كفاية حجم العينة: حيث يشـير معامل كايزر–ماير–أولكين (KMO) إلى كفاية حجم العينة، فقد كانت قيمة معامل كايزر–ماير–أولكين للصورتين على التوالي 0.98، 0.99، وتُعَدّ قيمًا ممتازة؛ لأنها أعلى من الحدّ الأدنى المقبول لهذا الاختبار الإحصائي، وهو 0.5 (Kaiser, 1974).
- مصفوفة معاملات الارتباط ليست مصفوفة وحدة في اختبار Bartlett's Test of Sphericity؛ لأنه دالّ إحصـــائيًا عند مســـتوى الدلالة 0.001، وهو أقل من 0.05؛ ولذا تكون مصـــفوفتا معاملات الارتباط لصورتي الامتحان ليستا مصفوفتي وحدة (Pett et al.,2003).، وتشير هذه النتائج مجتمعة إلى صلاحية بيانات صورتي الامتحان لإجراء التحليل العاملي.

بعدها أُنْجِزَ التحليل العاملي الاستكشافي بطريقة تحليل المحور الأساسي Principal Axis Factoring (PAF) على فقرات صورتي الامتحان، حيث إن التحليل العاملي بطريقة (PAF) مصمم لاستخراج العوامل التي تفسّر التباين المشترك بين المتغيرات، فهو يعمل على الكشف عن العوامل الأساسية التي تؤثر في استجابة الأفراد على الفقرات؛ ولذا فهو يوفّر نتائج أكثر دقة وتوافقًا مع نظرية الاستجابة للفقرة (Field, 2013)، ويوضح الجدول (12) نتائج التحليل العاملي لصورتي الامتحان النهائي لمادة الرياضيات المتقدمة لدبلوم التعليم العام.

الجدول (12) نتائج التحليل العاملي لبيانات صورتي الامتحان

الصورة الأولى 2023/ 2024م				الصورة الثانية 2024/ 2025م			
العوامل	الجذر الكامن للعامل	نسبة التباين المفسر من العامل%	التباين المفسر التراكمي%	العوامل	الجذر الكامن للعامل	نسبة التباين المفسر من العامل%	التباين المفسر التراكمي%
1	10.19	37.75	37.75	1	12.36	42.62	42.62
2	1.51	5.59	43.34	2	1.42	4.90	47.52
3	1.05	3.90	47.24				

فبالنسبة للصورة الأولى 2023/ 2024م أشار التحليل العاملي في الجدول (12) إلى أن عدد العوامل في هذه الصورة ثلاثة عوامل، فقد أشار إلى أن قيمة الجذر الكامن للعامل الأول 10.19، ونسبة التباين المفسر من هذا العامل 37.75%، وهو يزيد بمقدار سبعة أضعاف على نسبة التباين المفسر من العامل الثاني.

وبالنسبة للصورة الثانية من الامتحان 2024/ 2025م أشار التحليل العاملي في الجدول (12) إلى أن عدد العوامل في هذه الصورة عاملان، فقد أشار إلى أن قيمة الجذر الكامن للعامل الأول 12.36، ونسبة التباين المفسر من هذا العامل 42.62%، وهو يزيد بمقدار عشرة أضعاف على نسبة التباين المفسر من العامل الثاني، وهذا يدل على أن صورتي الامتحان تقيس كلتاهما بُعدًا واحدًا، وذلك بالاعتماد على محكّ أحادية البعد الذي يَعدّ الحدّ الأدنى للجذر الكامن أكبر من 2، ويكون عدد أفراد العينة كبيرًا جدًّا (Eigenvalue >2) (Jolliffe, 2002)، ويوضح الجدول (13) تشبعات فقرات صورتي امتحان مادة الرياضيات المتقدمة على العامل السائد لكل منها.

الجدول (13) تشبعات فقرات صورتي الامتحان على العامل السائد لكل صورة

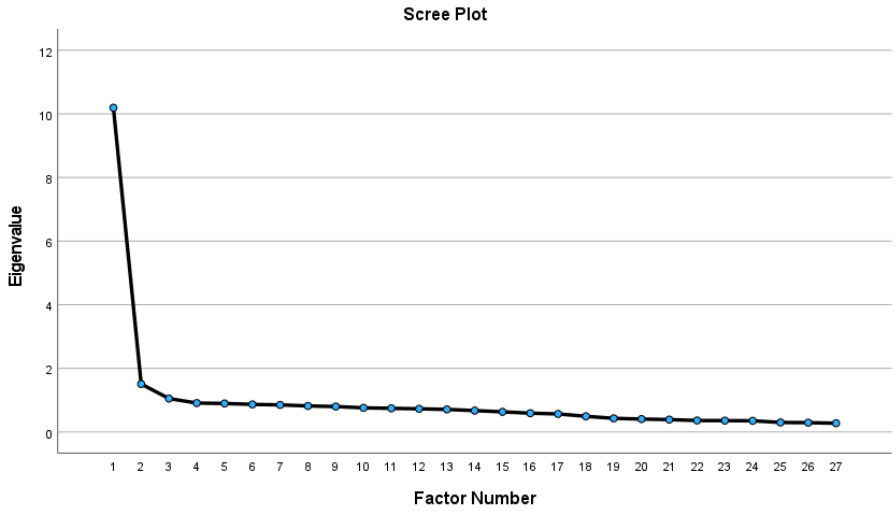
تشبعات الفقرات على العامل في الصورة الأولى			تشبعات الفقرات على العامل في الصورة الثانية		
الفقرة		التشبعات على العامل	الفقرة		التشبعات على العامل
النوع	رقم الفقرة	1	النوع	رقم الفقرة	1
اختيار من متعدد	1	0.33	اختيار من متعدد	1	0.22
	2	0.35		2	0.51
	3	0.41		3	0.50
	4	0.51		4	0.42
	5	0.36		5	0.51
	6	0.52		6	0.43
	7	0.38		7	0.58
	8	0.47		8	0.39
	9	0.30		9	0.66
	10	0.44		10	0.54
	11	0.54		11	0.48
	12	0.32		12	0.46
مقالي	1	0.61	مقالي	1	0.66
	2	0.83		2	0.71
	3	0.54		3	0.66
	4	0.64		4	0.67
	5	0.83		5	0.81
	6	0.73		6	0.76
	7	0.61		7	0.83
	8	0.59		8	0.71
	9	0.80		9	0.67
	10	0.83		10	0.72
	11	0.79		11	0.80
	12	0.78		12	0.82
	13	0.82		13	0.67
	14	0.76		14	0.83
	15	0.78		15	0.80
				16	0.79
				17	0.77

ومن جدول تشبعات الفقرات على العوامل (الجدول 13) نجد أن قيم تشبعات جميع فقرات امتحان الصورة الأولى 2023/ 2024م على البعد الأول كانت أكبر من 0.30، حيث تراوحت بين 0.32 و0.83، بينما كانت 3 فقرات منها تشبعاتها أكبر من 0.30 على البُعد الثاني، وأيضًا كانت تشبعات 3 فقرات منها أكبر من 0.30 على البُعد الثالث، وهذا يدلّ على أن هذه الفقرات أكثر ارتباطًا بالبُعد الأول، حيث يمكن الاستغناء عن البُعدين: الثاني والثالث؛ ولذا يمكن حسبان أن امتحان الصورة الأولى أحادي البُعد؛ لأن جميع الفقرات تتشبع على بُعد واحد، وقد أشار Harlow (2005) إلى أن قيم محك التشبع الجوهري أكبر من |0.30∓|.

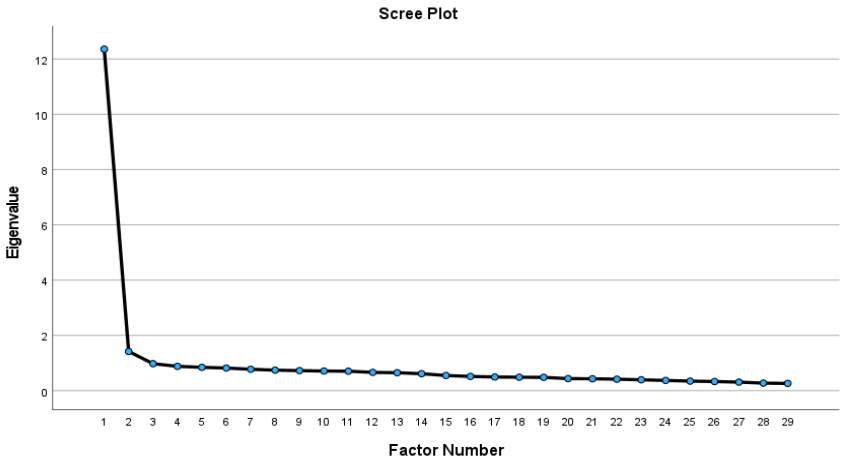
ونلاحظ من الجدول (13) كذلك أن قيم تشبعات جميع فقرات امتحان الصورة الثانية 2024/ 2025م على البعد الأول كانت أكبر من 0.30، فقد تراوحت بين 0.39 و0.83 ماعدا فقرة واحدة كان تشبعها 0.22، وكان أيضًا تشبعها على البعد الثاني سالبًا؛ فهي أقرب للبُعد الأول، بينما كانت 3 فقرات منها تشبعاتها أكبر من 0.30 على البُعد الثاني، وهذا يدلّ على أن هذه الفقرات أكثر ارتباطًا بالبُعد الأول، حيث يمكن الاستغناء عن البُعد الثاني؛ ولذا يمكن حسبان أن امتحان الصورة الثانية أحادي البُعد؛ لأن جميع الفقرات تتشبع على بُعد واحد.

وكذلك فإن Scree plot للتحليل العاملي لصورتي امتحان مادة الرياضيات المتقدمة لدبلوم التعليم العام شكل (14)، وشكل (15) يشيران إلى وجود عامل واحد تتجمع عليه فقرات الصورة الأولى 2024/2023م والصورة الثانية 2024/ 2025م على التوالي، حيث يوجد انحدار كبير بين العاملين: الأول والثاني في التمثيل البياني لقيم الجذور الكامنة للعوامل، وهو ما أشار إليه Hayes (2012).

الشكل (14) Scree plot للتحليل العاملي لبيانات الصورة الأولى



الشكل (15) Scree plot للتحليل العاملي لبيانات الصورة الثانية



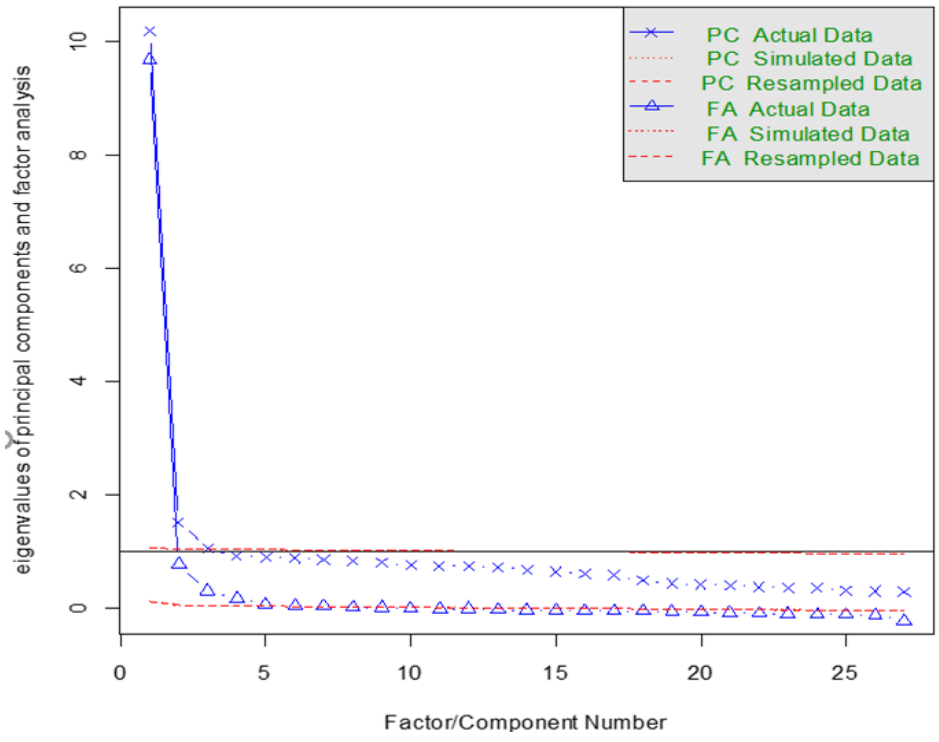
واسْتُعْمِلَ أيضًا التحليل العاملي المتوازي (Parallel Analysis) لتحديد العدد الأمثل للعوامل القابلة للاحتفاظ بها، حيث يُعَدّ التحليل العاملي المتوازي من أكثر الأساليب دقة في تحديد عدد العوامل، وموصى به على نطاق واسع في الأدبيات الإحصائية خصوصًا عند مقارنته بقاعدة Kaiser التقليدية، وذلك من خلال مقارنة الجذور الكامنة (Eigenvalues) الناتجة عن البيانات الفعلية مع تلك الناتجة عن بيانات عشوائية مماثلة في الحجم (Hayton et al., 2004; Horn, 1965).

ففي امتحان الصورة الأولى 2023/ 2024م أظهرت النتائج أن الجذور الكامنة للبيانات الفعلية تجاوزت نظيراتها العشوائية حتى العامل الثاني (أو الثالث)، مما يشير إلى أن عدد العوامل المناسبة في هذه الصورة هو عامل واحد.

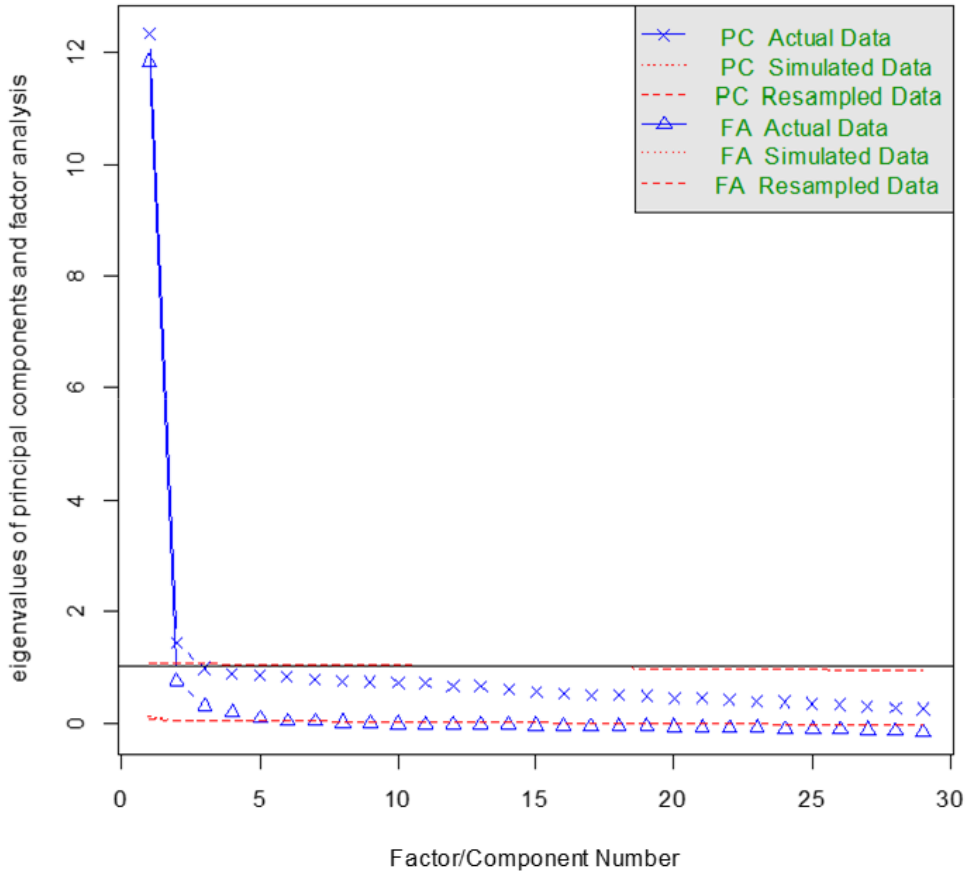
وفي الصورة الثانية 2024/ 2025م أظهرت النتائج أن الجذور الكامنة للبيانات الفعلية تجاوزت نظيراتها العشوائية، مما يشير إلى أن عدد العوامل المناسبة في هذه الصورة هو عامل واحد.

ويوضح الشكلان: (16، و17) نتائج التحليل العاملي المتوازي سواء لتحليل المكونات الأساسية (PC) في شكل ×، أو طريقة التحليل العاملي (FA) في شكل △ للصورتين على التوالي، حيث أظهر بدقة أحادية البُعد لصورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام.

الشكل (16) Scree plot للتحليل العاملي المتوازي لبيانات الصورة الأولى



الشكل (17) Scree plot للتحليل العاملي المتوازي لبيانات الصورة الثانية



مما سبق نستنتج أن صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام للعام 2023/ 2024م والعام 2024/ 2025م تتمتع بخاصية أحادية البُعد؛ ولذا فهي تمتاز بصدق البناء، وتتفق نتائج التحليل العاملي في هذه الدراسة مع دراسات كلّ من: (الزعابية وحسن، 2018؛ الشريفين، 2009، الشريفين والكيلاني، 2003؛ عبد الوهاب، 2015) من حيث تحقق أحادية البُعد، لكنها تختلف مع دراسات كلّ من: (الحبسي، 2018؛ الرواجفة والطراونة، 2016؛ الشريفين، 2009) من حيث عدد الأبعاد التي توزعت عليها فقرات الصور الامتحانية.

وأمكن كذلك تقدير الثبات، وذلك بحساب معامل الاتساق الداخلي باستعمال معامل ألفا كرونباخ ومعامل ماكدونالد أوميجا لصورتي الامتحان ويوضح الجدول (14) قيم هذه المعاملات لصورتي امتحان مادة الرياضيات المتقدمة.

الجدول (14) قيم معاملات ألفا كرونباخ وماكدونالد أوميجا

الصورة	عدد الفقرات	معامل ألفا كرونباخ	معامل ماكدونالد أوميجا
الصورة الأولى 2023/ 2024م	27	0.93	0.95
الصورة الثانية 2024/ 2025م	29	0.94	0.96

ويلاحظ من القيم في الجدول (14) أن قيم معامل ألفا كرونباخ كانت (0.93، 0.94) لصورتي الامتحان على التوالي، وكانت قيم معامل ماكدونالد أوميجا (0.95، 0.96) للصورتين على التوالي، وهي معاملات ثبات مرتفعة؛ حيث أنها تشير إلى تمتع صورتي الامتحان بثبات مرتفع Dunn et. al, (2014 ، وتتفق نتائج هذه الدراسة مع دراستي كلّ من: (الشريفين، 2009؛ الرواجفة والطراونة، 2016) كون أن الصور الاختبارية تتمتع بمعاملات ثبات مرتفعة، لكنها تختلف مع دراسة عبد الوهاب (2015) في مقدار الفرق بين معاملات الثبات في الصورتين.

النتائج المتعلقة بالسؤال الثاني ومناقشتها

للإجابة عن السؤال الثاني الذي ينصّ على "ما مدى مطابقة درجات صورتي الامتحان النهائي للفصل الدراسي الأول لمادة الرياضيات المتقدمة لدبلوم التعليم العام لافتراضات نظرية الاستجابة للفقرة؟" فقد تم التحقق من مطابقة البيانات لافتراضات نظرية الاستجابة للفقرة، وذلك كالآتي:

1. افتراض أحادية البُعد: تم التحقق من هذا الافتراض بالتحليل العاملي والتحليل العاملي المتوازي، وقد أشارت النتائج إلى أن صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام للعامين الدراسيين: 2023/ 2024م، و2024/ 2025م تحققان أحادية البُعد، حيث نوقِشَت النتائج المتعلقة بأحادية البُعد عند الإجابة عن السؤال الأول.

2. افتراض الاستقلال الموضعي: تم التحقق من افتراض الاستقلال الموضعي باستعمال حزمة (mirt)، وذلك عن طريق فحص مصفوفة ارتباطات البواقي المعيارية (standardized residual correlations) لصورتي الامتحان، وقد أظهرت النتائج أن مدى قيم معاملات الارتباط بين درجات البواقي للفقرات تتراوح (من −0.122 إلى 0.133) بالنسبة للصورة الأولى، وكان المدى لقيم معاملات ارتباطات البواقي المعيارية للصورة الثانية (من −0.110 إلى 0.126)، ويوضح الجدول (15) الإحصائيات المتعلقة بفحص الاستقلال الموضعي لصورتي الامتحان.

الجدول (15) ارتباطات البواقي المعيارية لفقرات صورتي الامتحان

الصورة	أقل ارتباط	الربيع الأول	الوسيط	المتوسط	الربيع الثالث	أعلى ارتباط
الأولى 2023 / 2024م	−0.122	−0.023	0.023	0.013	0.042	0.131
الثانية 2024/ 2025م	−0.110	−0.042	−0.018	−0.001	0.040	0.126

وتشير القيم في الجدول (15) من قيمة متوسط ارتباطات البواقي إلى أن الأخطاء موزعة بشكل طبيعي حول الصفر، وتشير أيضًا إلى تحقّق افتراض الاستقلال الموضعي لصورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام لعامي 2023/ 2024م و2024/ 2025م؛ لأنه يجب أن يكون أقل قيمة لارتباطات البواقي المعيارية ±0.30 كحد أقصى (Linacer, 2024; Yen, 1984).

مما سبق يتضح أن صورتي امتحان مادة الرياضيات المتقدمة لدبلوم التعليم العام لعامي 2023/ 2024م و2024/ 2025م مطابقة لافتراضات نظرية الاستجابة للفقرة، وتتفق هذه النتيجة مع نتيجة دراسات كلّ من: (الزعابية وحسن، 2018؛ القرني، 2023؛ الرواجفة والطراونة، 2016؛ عبد الوهاب، 2015).

النتائج المتعلقة بالسؤال الثالث ومناقشتها

للإجابة عن السؤال الثالث الذي ينصّ على "ما مدى مطابقة فقرات صورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة لنموذج نظرية الاستجابة للفقرة (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية)؟" فقد تم التحقق من مطابقة فقرات صورتي الامتحان النهائي لمادة الرياضيات المتقدمة لنماذج نظرية الاستجابة للفقرة (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية)، وحُلّلت الصورتان معًا باستعمال حزمة (mirt)، ويوضح الجدول (16) مؤشرات جودة المطابقة للفقرات ثنائية التصحيح لصورتي الامتحان للنموذج ثلاثي المعلمة (3PLM).

الجدول (16) مؤشرات جودة مطابقة الفقرات ثنائية التصحيح في صورتي الامتحان للنموذج الثلاثي

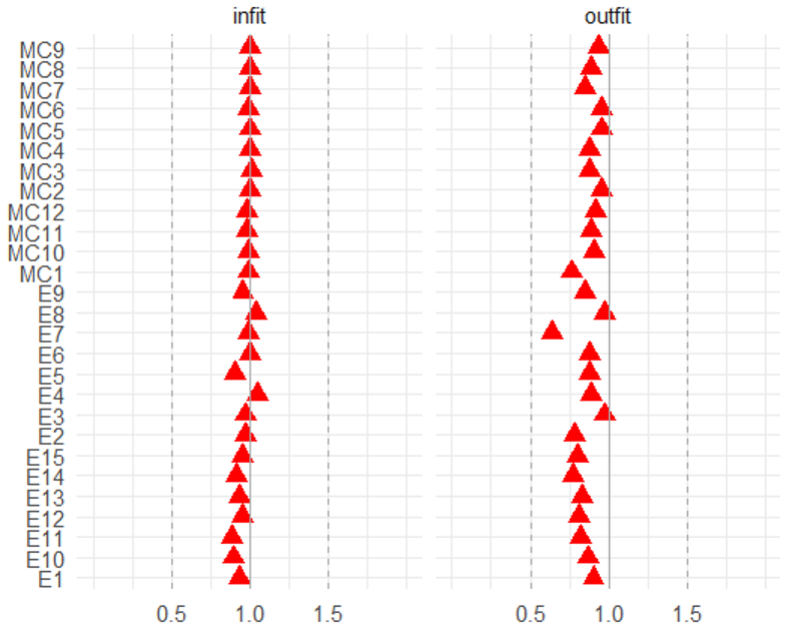
مؤشرات المطابقة	الصورة الأولى 2023/ 2024م	الصورة الثانية 2024/ 2025م
مؤشر M2(df)	3118.51(269)	6973.964(324)
مستوى الدلالة P–value	0.00	0.00
مؤشر الجذر التربيعي لمتوسط خطأ التقدير RMSEA	0.021	0.027
الجذر التربيعي لمتوسط البواقي المعيارية SRMSR	0.020	0.022
مؤشر المطابقة المعياري TLI	0.997	0.996
مؤشر المطابقة المقارن CFI	0.997	0.996

وقد قُيّمَت جودة مطابقة النموذج باستعمال اختبار M2 ومجموعة من مؤشرات المطابقة المناسبة لنماذج التحليل العاملي، وأظهرت النتائج في الجدول (16) أن قيمة مؤشر CFI بلغت 0.997، 0.996 للصورتين على التوالي، وقيمة مؤشر TLI بلغت 0.997، 0.996، وهي أعلى بكثير من الحدّ المقبول (≤ 0.95)؛ مما يدلّ على مطابقة ممتازة للنموذج، وكانت قيمة مؤشر RMSEA 0.021، 0.027 للصورتين على التوالي، وقيمة مؤشر SRMSR تساوي 0.020، 0.022 للصورتين على التوالي، وهي

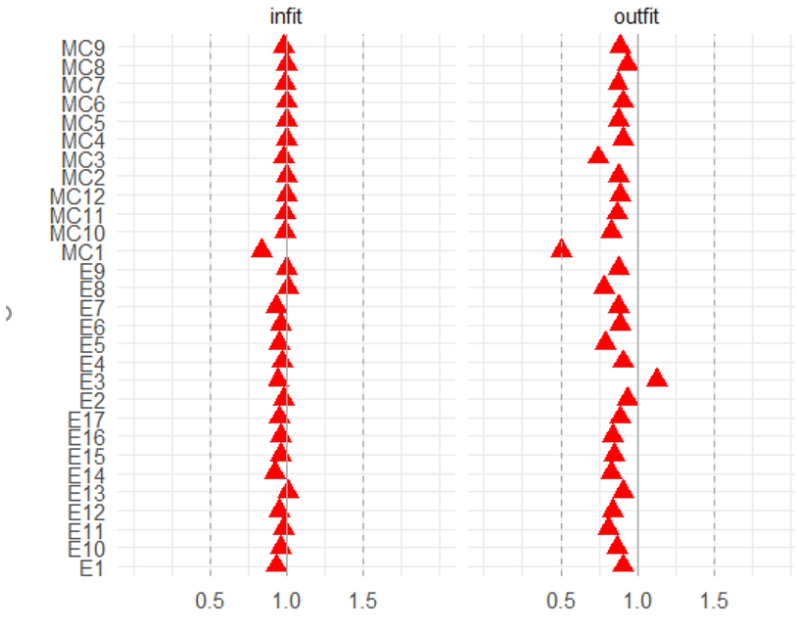
أقلّ من الحدود الحرجة المعتمدة (0.06 و0.08 على التوالي)؛ مما يعزّز من جودة المطابقة العامة (Hu & Bentler, 1999 ; Schumacker & Lomax, 2016)، وعلى الرغم من أن اختبار M2 كان دالًا إحصائيًّا ($p < 0.05$)، إلا أن مؤشرات المطابقة الثانية تؤكد ملاءمة البيانات في الصورتين للنموذج بشكل قويّ (Immekus et al., 2019).

واسْتُعْمِلَت أيضًا إحصائيات المطابقة الداخلية والخارجية Infit وOutfit ؛ لتقييم مدى مطابقة البيانات وملائمتها للنموذج، ويوضح الشكلان: (18، و19) أن جميع فقرات صورتي الامتحان تقع في النطاق المقبول إحصائيًا لـ Infit وOutfit (0.5 إلى 1.5) (Linacre, 2002)؛ مما يشير إلى اتساق الفقرات مع افتراضات النموذج، ويشير أيضًا إلى جودة بناء صورتي الامتحان وفعاليتهما في قياس البناء المصمم لقياسه، وكذلك فإنه لم يشِر إلى وجود فقرات متطرفة في الصورتين.

الشكل (18) يوضح إحصائيات Infit وOutfit لصورة الامتحان الأولى



الشكل (19) يوضح إحصائيات Infit وOutfit لصورة الامتحان الثانية



مما سبق يتضح أن فقرات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام لعامي 2023/ 2024م و2024/ 2025م مطابقة لنماذج نظرية الاستجابة للفقرة (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية)، وتتفق هذه النتيجة مع دراسة الزعابية وحسن (2018).

النتائج المتعلقة بالسؤال الرابع ومناقشتها

للإجابة عن السؤال الرابع الذي ينصّ على "ما تقدير معالم فقرات صورتي الامتحان النهائي للفصل الدراسي الأول في مادة الرياضيات المتقدمة باستعمال (النموذج ثلاثي المعالم (3PLM) لفقرات الاختيار من متعدد) و(نموذج التقدير الجزئي المعمم (G–PCM) للفقرات المقالية)؟" فقد قُدِّرَت معالم الفقرات باستعمال النموذج الثلاثي (3PLM) للفقرات ثنائية التصحيح ونموذج التقدير الجزئي المعمم (G–PCM) للفقرات متدرجة التصحيح لصورتي الامتحان، ويوضح الجدولان: (17، و18) معالم فقرات صورتي الامتحان على التوالي.

الجدول (17) معالم فقرات امتحان الصورة الأولى 2023/ 2024م

امتحان الصورة الأولى 2024/2023م						
الفقرات ثنائية التصحيح						
المعلمة	معلمة التمييز		معلمة الصعوبة		معلمة التخمين	
الفقرة	تقدير التمييز	خطأ التقدير	تقدير الصعوبة	خطأ التقدير	تقدير التخمين	خطأ التقدير
MC1	1.749	0.075	−2.580	0.145	0.071	0.137
MC2	1.434	0.048	−2.559	0.089	0.034	0.062
MC3	1.679	0.084	−1.390	0.112	0.313	0.058
MC4	2.046	0.088	−0.877	0.055	0.232	0.030
MC5	1.099	0.065	−1.389	0.201	0.181	0.087
MC6	1.986	0.053	−1.634	0.029	0.008	0.016
MC7	2.265	0.136	−0.442	0.054	0.498	0.019
MC8	2.480	0.124	−0.521	0.041	0.359	0.019
MC9	1.146	0.076	−0.976	0.176	0.360	0.058
MC10	1.840	0.081	−0.267	0.041	0.224	0.019
MC11	2.352	0.097	−0.010	0.021	0.118	0.011
MC12	2.341	0.172	0.542	0.025	0.333	0.010

الفقرات المقالية							
المعلمة		الصعوبة					
الفقرة	التمييز	عتبة الصعوبة 1	عتبة الصعوبة 2	عتبة الصعوبة 3	عتبة الصعوبة 4	عتبة الصعوبة 5	خطأ التقدير
E3	1.374	−1.953	−2.611	−1.641			0.065
E6	1.357	−1.331	−0.93	−0.651			0.104
E7	1.466	−1.153	−1.111	−2.534			0.062
E13	2.195	−1.333	−0.595	−0.301			0.062
E14	1.658	0.01	0.295	−0.514			0.068
E15	1.53	−0.059	−0.149	−1.093			0.07
E1	1.275	−1.057	−1.399	−2.384	−1.707		0.108
E4	0.876	−1.05	−0.625	−1.984	−1.966		0.051
E8	0.663	−2.142	0.99	−2.532	−1.143		0.075
E9	1.444	−0.619	−1.202	−0.175	−0.505		0.079
E10	1.681	−1.147	−0.481	−0.66	−0.469		0.088

	امتحان الصورة الأولى 2024/2023م						
	الفقرات ثنائية التصحيح						
	معلمة التخمين		معلمة الصعوبة		معلمة التمييز		المعلمة
	خطأ التقدير	تقدير التخمين	خطأ التقدير	تقدير الصعوبة	خطأ التقدير	تقدير التمييز	الفقرة
0.121	−1.279	0.384	−0.72	−0.316	−1.198	1.488	E2
0.086	−0.256	−0.005	−0.902	−0.921	−0.974	1.42	E5
0.116	−0.263	1.81	−0.361	−1.072	−0.437	1.332	E11
0.117	−2.834	−1.542	−0.281	0.308	0.163	0.959	E12

الجدول (18) معالم فقرات امتحان الصورة الثانية 2023/ 2024م

	امتحان الصورة الثانية 2024/ 2025م						
	الفقرات ثنائية التصحيح						
	معلمة التخمين		معلمة الصعوبة		معلمة التمييز		المعلمة
	خطأ التقدير	تقدير التخمين	خطأ التقدير	تقدير الصعوبة	خطأ التقدير	تقدير التمييز	الفقرة
	0.137	0.071	0.145	−2.580	0.075	1.749	MC1
	0.062	0.034	0.089	−2.559	0.048	1.434	MC2
	0.058	0.313	0.112	−1.390	0.084	1.679	MC3
	0.030	0.232	0.055	−0.877	0.088	2.046	MC4
	0.087	0.181	0.201	−1.389	0.065	1.099	MC5
	0.016	0.008	0.029	−1.634	0.053	1.986	MC6
	0.019	0.498	0.054	−0.442	0.136	2.265	MC7
	0.019	0.359	0.041	−0.521	0.124	2.480	MC8
	0.058	0.360	0.176	−0.976	0.076	1.146	MC9
	0.019	0.224	0.041	−0.267	0.081	1.840	MC10
	0.011	0.118	0.021	−0.010	0.097	2.352	MC11
	0.010	0.333	0.025	0.542	0.172	2.341	MC12
	الفقرات المقالية						
	الصعوبة						المعلمة
خطأ التقدير	عتبة الصعوبة 5	عتبة الصعوبة 4	عتبة الصعوبة 3	عتبة الصعوبة 2	عتبة الصعوبة 1	التمييز	الفقرة
0.073				−0.647	−1.131	1.48	E4

0.048				0.052	−0.963	2.152	E6
0.064			−1.54	−1.532	−0.67	1.405	E1
0.053			−0.783	0.044	−0.328	1.766	E5
0.08			−1.293	−0.406	−2.12	1.505	E8
0.065			−1.004	−1.168	−1.539	1.366	E9
0.073			−1.045	−0.977	−1.636	1.676	E10
0.061			−0.609	−1.919	−1.774	1.684	E13
0.057			−0.815	−0.355	−0.304	1.638	E16
0.066			0.445	−0.845	−0.748	1.697	E17
0.08		−0.532	−0.273	−1.411	−1.897	1.124	E2
0.072		−2.19	−0.86	−0.722	−0.952	0.923	E3
0.059		−0.503	0.047	−0.744	−0.407	1.555	E7
0.075		−1.458	−0.167	−0.761	0.539	1.219	E11
0.067		−0.704	0.288	−0.401	−0.23	1.351	E15
0.076	−0.753	−0.354	−0.92	−0.796	−1.322	1.558	E12
0.058	0.086	−0.579	−0.049	0.011	−0.291	1.443	E14

وتشير النتائج في الجدول (17) أن معلمة الصعوبة للفقرات ثنائية التصحيح لامتحان الصورة الأولى 2023/ 2024م تراوحت بين (−2.580 إلى 0.542) وبخطأ قياس تراوح بين (0.021 إلى 0.201)؛ مما يدلّ على أن الفقرات ثنائية التصحيح في هذه الصورة تراوحت صعوبتها بين السهلة والمتوسطة، حيث يشير Baker (2001) إلى أن القيم المقبولة لمعلمة الصعوبة تقع في المدى(−2 إلى 2)، بينما تراوحت قيمة معامل التمييز لها بين (1.099 إلى 2.480) وبخطأ قياس تراوح بين (0.048 إلى 0.172)؛ مما يدلّ على أن فقرات الصورة الأولى تمتعت بتمييز عالٍ إلى تمييز عالٍ جدًّا، حيث يصنف Baker (2001) التمييز إلى a<0.20 أنه تمييز منخفض جدًّا، والقيم $0.21 < a < 0.40$ أنه تمييز منخفض، والقيم بين $0.41 < a < 0.80$ أنه تمييز متوسط، والقيم $0.81 < a < 1.00$ أنه تمييز مرتفع، والقيم $a > 1.00$ تمييز عال جدًّا.

وأظهرت النتائج أيضًا أن قيمة التخمين لهذه الفقرات تراوحت بين (0.008 إلى 0.498) وبخطأ قياس تراوح بين (0.010 إلى 0.137)، وهذا يدلّ على أن مستوى التخمين في الفقرات ثنائية الاستجابة

لامتحان الصورة الأولى تراوح بين القيم القريبة من الصفر– وهذا يعني أن احتمال حلّها بالتخمين ضعيف– وبين قيم قريبة من 50%، أي أن احتمال حلّها بالتخمين كبير، حيث تشير De Ayala (2009) إلى أن قيم معلمة التخمين تتراوح بين (0 إلى 1)، وأفضل قيم للتخمين كلما اقتربت من الصفر؛ لأنها تدلّ على احتمالية حلّها بالتخمين أقلّ.

بينما قُدِّرَت معالم فقرات الأسئلة المقالية الصعوبة والتمييز لصورتي الامتحان الصورة الأولى 2023/ 2024م والصورة الثانية 2025/2024م باستعمال نموذج التقدير الجزئي المعمم (G–PCM)، وأظهرت النتائج في الجدول (17) أن قيم التمييز للفقرات المقالية للصورة الأولى تراوحت بين (0.959 إلى 1.681)، وهي تعبِّر عن تمييز تراوح من تمييز عالٍ إلى تمييز عالٍ جدًّا (Baker, 2001) ، وتراوحت عتبات الصعوبة للفقرات المقالية للصورة الأولى من المتوسطة إلى السهلة جدًّا (Baker, 2001).

وتشير النتائج في الجدول (18) إلى أن معلمة الصعوبة للفقرات ثنائية التصحيح لامتحان الصورة الثانية 2024/ 2025م تراوحت بين (–3.142 إلى0.198) وبخطأ قياس تراوح بين (0.016 إلى 0.354)؛ مما يدلّ على أن الفقرات ثنائية التصحيح في هذه الصورة تراوحت صعوبتها بين السهلة والمتوسطة، حيث يشير Baker (2001) إلى أن القيم المقبولة لمعلمة الصعوبة تقع في المدى(–2 إلى 2)، بينما تراوحت قيمة معامل التمييز لها بين (1.458 إلى 3.188) وبخطأ قياس تراوح بين (0.067 إلى 0.168)؛ مما يدلّ على أن فقرات الصورة الثانية تمتعت بتمييز عالٍ جدًّا، حيث يصنّف Baker (2001) التمييز إلى: $a<0.20$ أنه تمييز منخفض جدًّا، والقيم $0.21 < a < 0.40$ أنه تمييز منخفض، والقيم بين $0.41 < a < 0.80$ أنه تمييز متوسط، والقيم $0.81 < a < 1.00$ أنه تمييز مرتفع، والقيم $a > 1.00$ تمييز عال جدًّا، وأظهرت النتائج كذلك أن قيمة التخمين لهذه الفقرات تراوحت بين (0.0239 إلى 0.402) وبخطأ قياس تراوح بين (0.008 إلى 0.393)، وهذا يدلّ على أن مستوى التخمين في الفقرات ثنائية الاستجابة لامتحان الصورة الثانية تراوحت بين القيم القريبة من

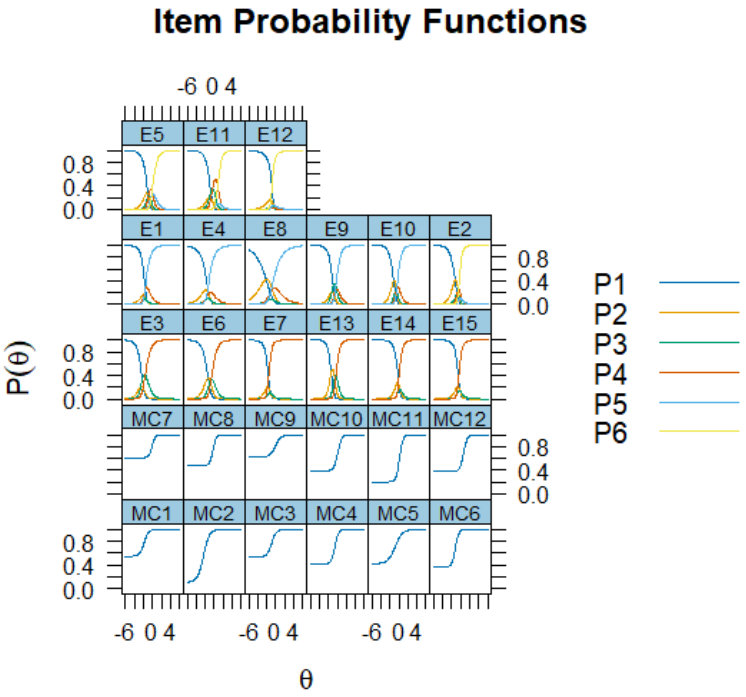
الصفر– وهذا يعني أن احتمال حلّها بالتخمين ضعيف– وبين قيم قريبة من 40%، أي أن احتمال حلّها بالتخمين كبير، حيث تشير De Ayala (2009) إلى أن قيم معلمة التخمين تتراوح بين (0 إلى 1)، وأفضل قيم للتخمين كلما اقتربت من الصفر؛ لأنها تدلّ على أن احتمالية حلّها بالتخمين أقل.

وتشير النتائج في الجدول (18) أيضًا إلى معالم الفقرات المقالية الصعوبة والتمييز لامتحان الصورة الثانية، وقد أمكن تقديرها باستعمال نموذج التقدير الجزئي المعمم (G–PCM)، حيث أظهرت النتائج في الجدول (18) أن قيم التمييز للفقرات المقالية للصورة الثانية تراوحت بين (0.923 إلى 2.152)، وتعبّر عن تمييز تراوح من تمييز عالٍ إلى تمييز عالٍ جدًّا (Baker, 2001)، وكذلك تراوحت عتبات الصعوبة للفقرات المقالية للصورة الثانية من المتوسطة إلى السهلة جدًّا (Baker, 2001).

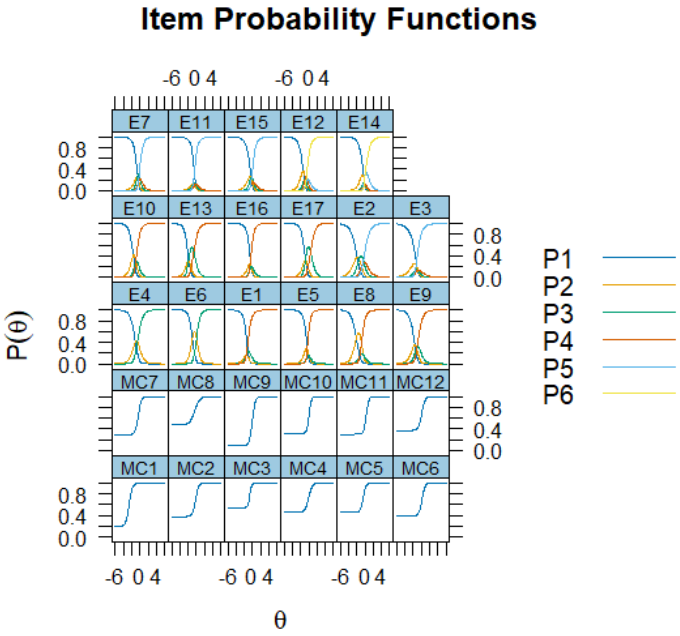
ويتضح من معالم فقرات صورتي الامتحان أن الصورتين غير متكافئتين، وهذا يؤكد ضرورة إجراء معادلة لصورتي الامتحان، وتتفق نتائج هذه الدراسة مع دراستي (الشريفين، 2009؛ العبيدي، 2018)؛ لأن فقرات الصور الامتحانية تتمتع بمعالم مقبولة تقع في المنتصف.

ومن منحنى خصائص الفقرة لفقرات امتحان الصورة الأولى والصورة الثانية (الشكلان: 20، و21 على التوالي)؛ حيث أن منحنى خصائص الفقرة (ICC) يربط بين قدرة المفحوص واحتمال إجابته على الفقرة، ويتأثر منحنى خصائص الفقرة بتمييز الفقرة وكذلك في الفقرات الثنائية يتأثر بمقدار التخمين فيها، ويوضح الرسم أن الفقرات الثنائية لها منحنى واحد فقط يتغير شكله باختلاف خصائص الفقرة بينما في الفقرات المتدرجة يكون منحنى خصائص الفقرة ممثَّلًا بأكثر من منحنى، فكلّ عتبة من عتبات الصعوبة يمثّلها منحنى مختلف عن منحنيات العتبات الأخرى، وذلك حسب صعوبة العتبة؛ لأن التمييز ثابت على مستوى الفقرة (Embretson & Reise, 2000).

الشكل (20) منحنى خصائص الفقرة لفقرات امتحان الصورة الأولى 2023/ 2024م

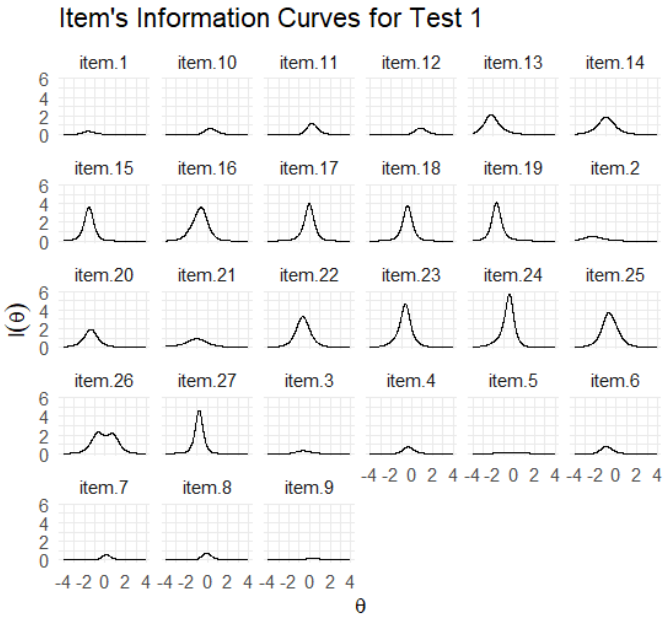


الشكل (21) منحنى خصائص الفقرة لفقرات امتحان الصورة الثانية 2024/ 2025م

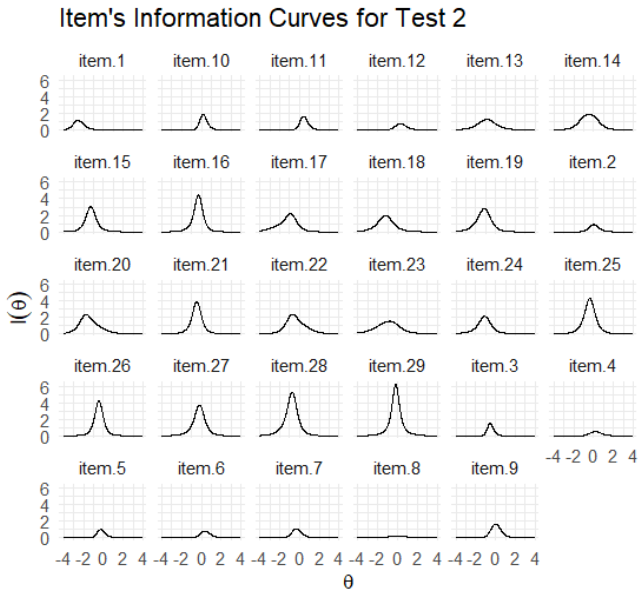


ويوضح الشكلان: (22، و23) منحنى دالة معلومات فقرات الامتحان للصورتين: الأولى والثانية على التوالي، والذي يلخّص مستوى فاعلية الفقرات ودقّة قياسها للقدرة مع الأخذ في الحسبان جميع خصائص الفقرات الصعوبة والتمييز والتخمين (De Ayala, 2009).

الشكل (22) منحنيات دالة معلومات فقرات الصورة الأولى 2023/ 2024م



الشكل (23) منحنيات دالة معلومات فقرات الصورة الثانية 2024/ 2025م

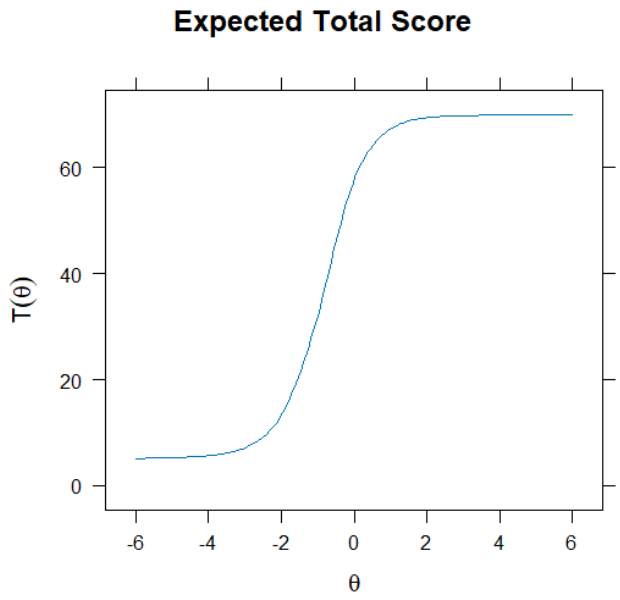


ويتضح من الشكل (22) أن معظم فقرات الصورة الأولى تعطي أكبر دالة معلومات في المنتصف، أي أن الفقرات صُمِّمَت؛ لتلائم الطلبة ذوي القدرات المتوسطة ماعدا الفقرات: 1، و2، و3، و5، و9 التي كانت دالة المعلومات لها منخفضة.

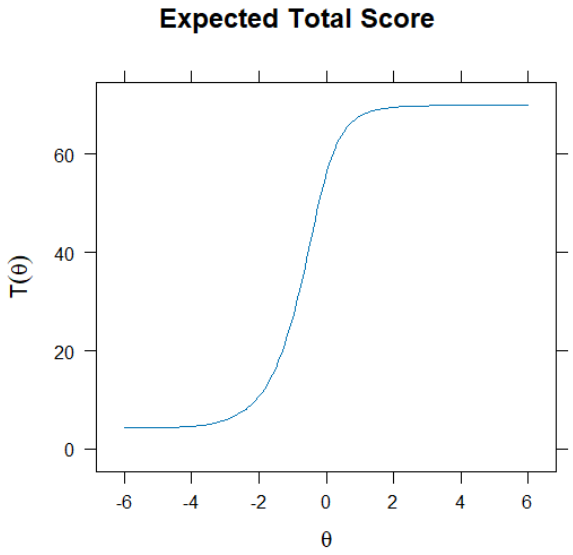
ويتضح من الشكل (23) أن معظم فقرات الصورة الثانية تعطي أكبر دالة معلومات في المنتصف، أي أن الفقرات صُمِّمَت؛ لتلائم الطلبة ذوي القدرات المتوسطة ماعدا الفقرتين: 4، و8، فقد كانت دالة المعلومات لها منخفضة.

ويوضح منحنى الدرجة الكلية المتوقعة (Expected Total Score Curve) لصورتي الامتحان (الشكلان: 24، و25) الذي يربط بين القدرة والدرجة المتوقعة، ونلاحظ أن شكل المنحنى للصورتين على شكل حرف S، وهو الشكل الطبيعي في نظرية الاستجابة للفقرة؛ مما يدلّ على أن صورتي الامتحان جيدتان لتمييز الطلبة من قدرات منخفضة إلى عالية، وأنهما فعّالتان تحديدًا في المدى [−2، 2]، وعند القدرات المنخفضة تكون الدرجات المتوقعة منخفضة أيضًا؛ مما يشير إلى أن الأفراد ذوي القدرات المتدنية يحققون درجات متدنية، ونلاحظ أيضًا استقرار المنحنيين عند القدرات العليا؛ مما يشير إلى تشبُّع الأداء لدى الأفراد ذوي القدرات العليا، حيث إن العلاقة بين القدرة الكامنة والدرجة الكلية تأخذ شكلًا لوجستيًا؛ مما يعكس زيادة حادّة في الدرجات ضمن نطاق معيّن من القدرة وتشبّعًا في الدرجات عند الأطراف (Embretson & Reise, 2000).

الشكل (24) منحنى الدرجة الكلية المتوقعة لصورة الامتحان الأولى 2023/ 2024م

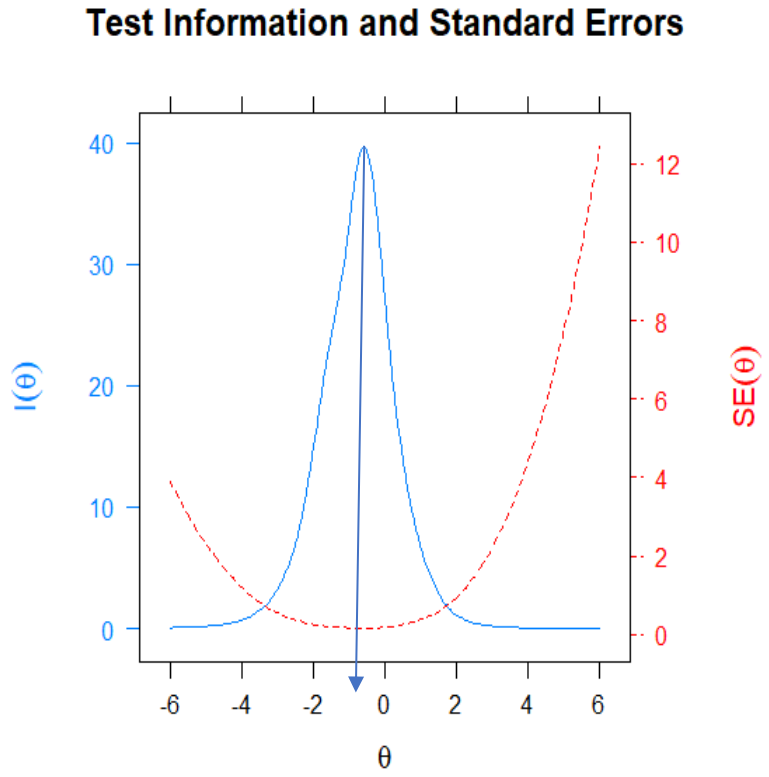


الشكل (25) منحنى الدرجة الكلية المتوقعة لصورة الامتحان الثانية 2024/ 2025م

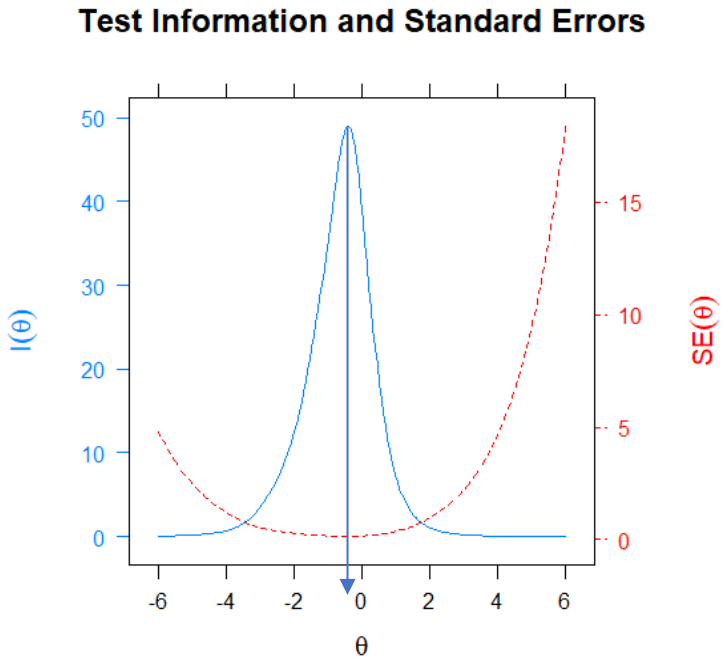


ويوضّح منحنى دالة معلومات الامتحان Test Information Function (TIF) لصورتي الامتحان (الشكلان: 26، و27 على التوالي) الذي يقدم مصدرين أو مؤشرين لدقّة القياس، وهما حجم المعلومات ممثّلًا بالخطّ المتصل والخطأ المعياري ممثّلًا بالخطّ المتقطّع، أن صورتي الامتحان تعطي أكبر قدر من المعلومات في المنتصف، ويكون لها أقلّ خطأ معياري للتقدير، بمعنى أن صورتي الامتحان تمتازان بدقّة عالية جدًّا في المنتصف، وتقلّ هذه الدقّة كثيرًا في القدرات المنخفضة والعالية، وهذا يؤكّد ملائمة صورتي الامتحان لطلبة القدرات المتوسطة، فقد أشار Bond & Fox (2015), and De Ayala (2009) إلى أن الشكل المثالي لدالة معلومات الاختبار يجب أن يُظهر أكبر قدر من المعلومات في المدى المتوسط من القدرة، ويكون الخطأ المعياري أعلى في الأطراف على اليمين واليسار، وهذا ينطبق على دالة معلومات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام، وهذا الأمر يتفق مع دراستي (Istiyono et al., 2020; Martin & Lazendic, 2018).

الشكل (26) دالة معلومات الاختبار لصورة الامتحان الأولى 2023/ 2024م



الشكل (27) دالة معلومات الاختبار لصورة الامتحان الثانية 2023/ 2024م



النتائج المتعلقة بالسؤال الخامس ومناقشتها

للإجابة عن السؤال الخامس الذي ينصّ على "ما مدى مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة وتقدير قدرات الأفراد؟" فقد تم التحقق من مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة عن طريق التحقّق من مؤشّر مطابقة الأفراد لنماذج نظرية الاستجابة للفقرة Infit، حيث تراوحت قيمة مؤشّر المطابقة للأفراد على صورتي الامتحان الأولى، والثانية على التوالي (0.0898 إلى 5.621)، (0.055 إلى 6.340).

ويوضح الجدول (19) نسبة الأفراد المطابقين للنموذج لصورتي الامتحان، وذلك بالاعتماد على معيار القيمة المقبولة لمؤشّر المطابقة Infit، حيث إن القيمة المقبولة لمؤشر المطابقة Infit تقع في المدى [0.5 إلى 2] (Linacre, 2017).

الجدول (19) نسبة الأفراد المطابقين للنموذج بناءً على مؤشّر المطابقة Infit لصورتي الامتحان

المجموعة الأولى 2024/2023م					المجموعة الثانية 2025/ 2024م				
عدد أفراد العينة	المطابقين		غير مطابقين		عدد أفراد العينة	المطابقين		غير مطابقين	
	العدد	النسبة	العدد	النسبة		العدد	النسبة	العدد	النسبة
24619	21415	%87	3204	%13	28659	25684	90%	2975	10%

من النتائج في الجدول (19) نلاحظ أن نسبة الأفراد المطابقين لنموذج نظرية الاستجابة للفقرة بلغت 87% ، و90% لصورتي الامتحان على التوالي، وتُعَدّ هذه النسب مقبولة؛ لأن نسب مطابقة الأفراد المقبولة في التطبيقات التربوية يقع في الحدود بين 85% و90% (Embretson & Reise, 2000; Hambleton et al. 1991; Wilson, 2005)، ونلاحظ أيضا من الجدول (19) وجود نسبة من الأفراد غير مطابقين لنموذج نظرية الاستجابة للفقرة على الصورتين مع زيادة النسبة في الصورة الأولى، حيث تشير نسبة عدم المطابقة إلى وجود مجموعة من الطلبة أجابوا إجابات لا تتماشى مع مستواهم العام، فمثلًا أخطأوا في فقرات سهلة، ونجحوا في الإجابة عن فقرات أصعب، أو وجود

حالات من الملل أو الإهمال في الإجابة من بعض الطلبة (2009) De Ayala، ويشير Embretson & Reise (2000) إلى عدم وجود طريقة لتفسير أسباب عدم المطابقة أو تفسير لنسب المطابقة الضعيفة.

وتتفق هذه النتائج مع دراسة (الشريفين، 2009) في وجود مجموعة من الأفراد غير مطابقين لنماذج نظرية الاستجابة للفقرة.

وقُدِّرَت أيضًا قدرات الأفراد، حيث تراوحت قدرات الأفراد لصورتي الامتحان [−3.187 إلى 1.666]، و[−3.043 إلى 1.686] على التوالي، ويوضح الجدول (20) الإحصاءات الوصفية لتوزيع قدرات الأفراد.

الجدول (20) الإحصاءات الوصفية لتوزيع قدرات الأفراد

المجموعة	المجموعة الأولى 2023/ 2024م	المجموعة الثانية 2024/ 2025م
أعلى قيمة للقدرة	1.666	1.686
أقل قيمة للقدرة	−3.187	−3.043
متوسط قدرات الأفراد	−0.002	−0.002
الانحراف المعياري للقدرات	0.964	0.963
التفلطح	−0.30	−0.53
الالتواء	−0.32	−0.09

نلاحظ من الجدول (20) تساوي المتوسّط الحسابي والانحراف المعياري لقدرات الأفراد لصورتي الامتحان، وهذا يدلّ على أن مجموعتي الأفراد تمتلكان المستوى نفسه والمهارة نفسها في مادة الرياضيات المتقدمة، ويدلّ أيضًا تساوي الانحراف المعياري لقدرات مجموعتي الأفراد على أن توزيع القدرات داخل كلا المجموعتين متشابه، ومنه يمكن القول: إن مجموعتي الأفراد متكافئتان، ونلاحظ كذلك من قيم التفلطح والالتواء أن قيم القدرة لطلبة المجموعتين تتوزّع توزيعًا طبيعيًا، حيث إن قيم التفلطح والالتواء كانت ضمن النطاق الطبيعي لمستويي التفلطح والالتواء (−3 إلى +3) (George & Mallery, 2010).

الجدول (21) توزيع الأفراد على فئات القدرات لمجموعتي الأفراد

المجموعة		المجموعة الأولى 2023/ 2024م		المجموعة الثانية 2024/ 2025م	
	الفئة	العدد	النسبة المئوية	العدد	النسبة المئوية
توزيع قدرات الأفراد على فئات	$\theta \leq -3$	37	0.15%	23	0.08%
	$-2 \geq \theta > -3$	550	2.23%	472	1.65%
	$-1 \geq \theta > -2$	3370	13.69%	3946	13.77%
	$0 \geq \theta > -1$	7745	31.46%	10128	35.34%
	$1 \geq \theta > 0$	9077	36.87%	9089	31.71%
	$2 \geq \theta > 1$	3840	15.60%	5001	17.45%
	المجموع	24619	100.00%	28659	100.00%

يوضح الجدول (21) النسب المئوية لفئات القدرة للأفراد في المجموعتين، حيث يظهر أن توزيع قدرات مجموعتي الأفراد متقارب لغالب هذه الفئات، وبشكل عام تشير النتائج إلى تشابه توزيع قدرات الأفراد على صورتي الامتحان؛ مما يدلّ على أن مجموعتي الطلبة متكافئتان، حيث يشير Hambleton and Swaminathan (1985) إلى أهمية تماثل توزيع القدرات لضمان تكافؤ المجموعتين ، ويشير Kolen and Brennan (2014) إلى أن تشابه المتوسّط والانحراف المعياري لقدرات مجموعتين من الأفراد مؤشّر أساسي لتكافؤ المجموعتين، وهذا يؤكّد أن التصميم المناسب لمعادلة صورتي الامتحان النهائي في مادة الرياضيات المتقدمة لدبلوم التعليم العام هو تصميم المجموعات العشوائية أو المتكافئة.

النتائج المتعلقة بالسؤال السادس ومناقشتها

للإجابة عن السؤال السادس الذي ينصّ على "ما قيم الدرجات المتعادلة لصورتي الامتحان النهائي لمادة الرياضيات المتقدمة لدبلوم التعليم العام؟" فقد أُنْجِزَت عملية معادلة درجات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال برنامج R، وذلك بتوظيف تصميم المجموعات العشوائية المتكافئة مع استعمال طريقة المتوسّط / الانحراف المعياري (Mean/ Sigma) لتقدير معاملات الربط في عملية المعادلة التي تستعمل المتوسّط والانحراف المعياري لحساب التحويل الخطيّ، واعْتُمِدَ في

عملية المعادلة امتحان الصورة الأولى 2023/ 2024م كصورة مرجعية تُفَسَّر به درجات الصورة الثانية 2024/ 2025م، وقد أُنْجِزَت المعادلة بعدّة طرائق:

1. معادلة درجات صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستخدام الدرجات الخام.

لإيجاد الدرجات المتعادلة باستعمال الدرجات الخامّ فقد اعْتُمِدَت الصورة الأولى من امتحان مادة الرياضيات المتقدمة لدبلوم التعليم العام لعام 2023/ 2024م صورة مرجعية؛ ليُعْتَمَد عليها في تفسير درجات الصورة الثانية، فقد حُوِّلَت الدرجات الخامّ للمفحوصين على الصورة الثانية إلى ما يقابلها على الصورة الأولى، ويوضح الجدول (22) الدرجات الخامّ المتعادلة على صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام.

الجدول (22) الدرجات الخامّ على الصورة الأولى وما يقابلها على الصورة الثانية

الدرجة الخام على الصورة الأولى 2023/ 2024م	الدرجة الخام على الصورة الثانية 2024/ 2025م	الدرجة الخام على الصورة الأولى 2023/ 2024م	الدرجة الخام على الصورة الثانية 2024/ 2025م	الدرجة الخام على الصورة الأولى 2023/ 2024م	الدرجة الخام على الصورة الثانية 2024/ 2025م
0	0	24	19	48	43
1	1	25	20	49	44
2	2	26	21	50	46
3	3	27	22	51	47
4	3	28	23	52	48
5	4	29	24	53	49
6	5	30	25	54	51
7	6	31	26	55	52
8	7	32	27	56	54
9	8	33	28	57	55
10	8	34	29	58	56
11	9	35	30	59	58
12	10	36	31	60	59

الدرجة الخام على الصورة الأولى 2023/ 2024م	الدرجة الخام على الصورة الثانية 2024/ 2025م	الدرجة الخام على الصورة الأولى 2023/ 2024م	الدرجة الخام على الصورة الثانية 2024/ 2025م	الدرجة الخام على الصورة الأولى 2023/ 2024م	الدرجة الخام على الصورة الثانية 2024/ 2025م
13	10	37	31	61	60
14	11	38	32	62	62
15	12	39	33	63	63
16	13	40	34	64	64
17	14	41	35	65	65
18	14	42	36	66	66
19	15	43	37	67	67
20	16	44	38	68	68
21	17	45	40	69	69
22	17	46	41	70	70
23	18	47	42		

من الجدول (22) نلاحظ أن الدرجات الدنيا متقاربة بين الصورتين، ثم يبدأ الفارق بين الدرجات الخام للصورتين بالزيادة إلى أن تبدأ الدرجات مرة أخرى بالتقارب في الدرجات العشر العليا، ونلاحظ أيضًا من الجدول (22) أن امتحان الصورة الأولى كان أسهل من امتحان الصورة الثانية، حيث أن الدرجة على الصورة الثانية تقابل درجة أعلى منها على الصورة الأولى، وهذه النتيجة تتفق مع دراسة المحرزيّ (2014) في أن الصورة المرجعية أسهل من الصورة الثانية.

وقد حُسِبَت الإحصاءات الوصفية للدرجات المتعادلة على صورتي الامتحان باستعمال الدرجات الخامّ، ويُقدِّم الجدول (23) ملخصًا لتلك النتائج.

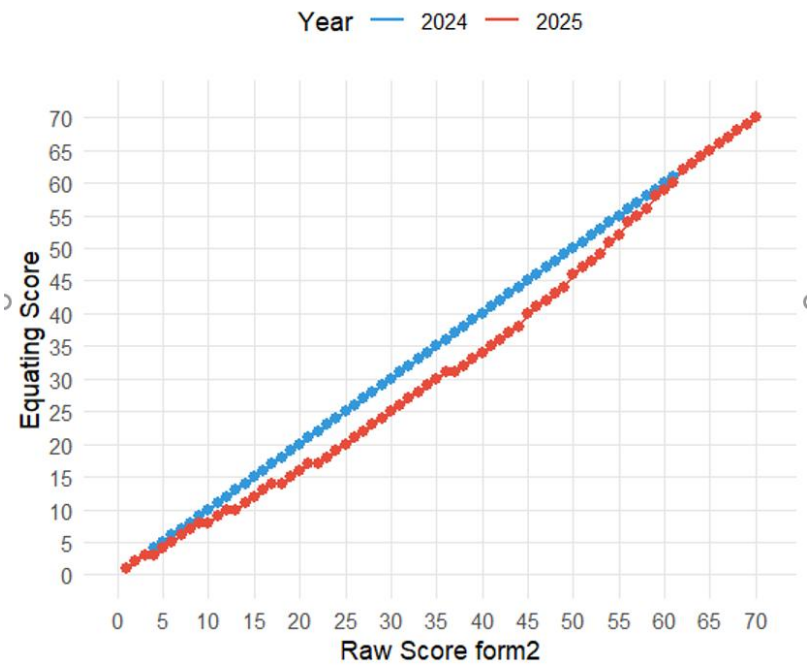
الجدول (23) الإحصاءات الوصفية للدرجات المتعادلة على صورتي الامتحان والفروق بينها

صورة الامتحان	المتوسط الحسابي	الانحراف المعياري	اختبار ت	مستوى الدلالة	حجم الأثر
الدرجات على الصورة الأولى (الدرجة الكلية)	35	20.64			
الدرجات المعادلة على الصورة الثانية	31.87	20.86	1.26	0.21	0.09
الفروق بين الدرجات المتعادلة على الصورتين	3.13	2.13			

يلاحظ من الجدول (23) اختلاف الإحصاءات الوصفية للدرجات الخام المتعادلة على الصورتين، فقد كانت الإحصاءات الوصفية للدرجات على الصورة الثانية المعادِلة للدرجات على الصورة الأولى (الدرجة الكلية) أقلّ من الإحصاءات الوصفية للصورة المرجعية (الأولى) مع التأكيد أن الفرق في الانحراف المعياري كان صغيرًا، وهذا يدلّ على أن تشتت الدرجات في الصورتين متقارب جدًّا، وكانت الدرجة الخام على الصورة الأولى تعادلها درجة خام أقل منها أو تساويها على الصورة الثانية، وأوضحت الفروق بين الدرجات الخام المتعادلة أن الفرق كان دائمًا لصالح الدرجات الخام للصورة الأولى؛ لأن الدرجات الخام على الصورة الأولى 2023/ 2024م كانت أعلى من الدرجات الخام المعادلة له على الصورة الثانية 2024/ 2025م، حيث يؤكد Kolen and Brennan (2014) تحقّق عملية المعادلة عند تقارب نتائج الإحصاءات الوصفية للدرجات المتعادلة لصورتي الامتحان باستعمال الدرجات الخام، وكذلك فإن اختبار t لعينتين مستقلتين يظهر أن متوسط الدرجات المعادِلة على الصورة الثانية لا يختلف عن متوسط المجموع الكلي 35، حيث أن مستوى الدلالة أعلى من 0.05، وبحجم تأثير 0.09، وهو حجم تأثير قليل، وهذا يؤكد تحقق عملية المعادلة.

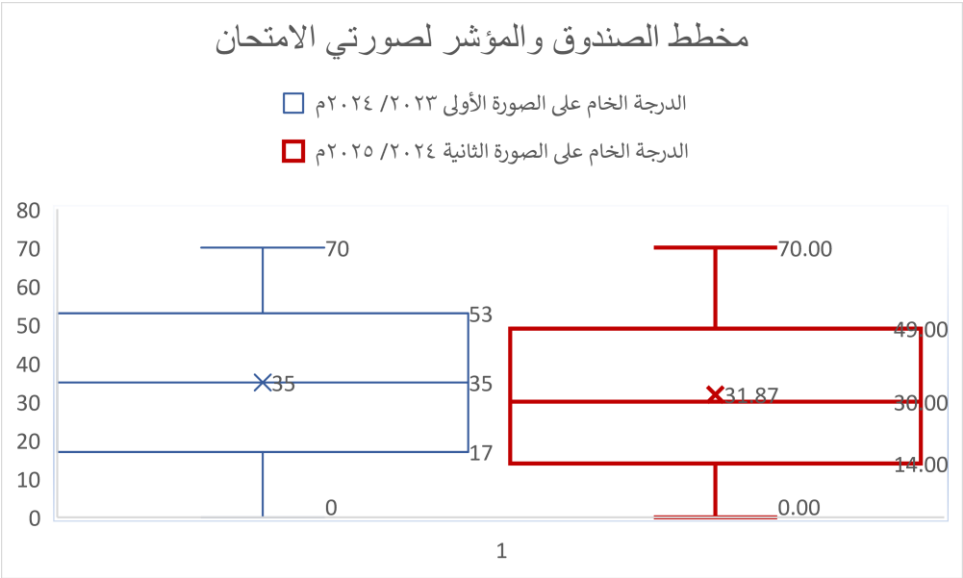
ويوضح الشكل (28) الدرجات الخام المعادلة للصورة الثانية مع ما يقابلها من الدرجات الخام للصورة المرجعية، حيث نلاحظ أن درجات الصورة الأولى 2023/ 2024م هي الأعلى دائمًا ما عدا في الأطراف، فإن الخطين منطبقان.

الشكل (28) الدرجات الخام المعادلة للصورة الثانية مع ما يقابلها من الدرجات الخام للصورة المرجعية



ويوضح الشكل (29) مخطط الصندوق والمؤشر لصورتي الامتحان، حيث نلاحظ أن الربيع الأدنى والربيع الأعلى والوسيط للدرجات المعادِلَة على الصورة الثانية كانت أقل من الدرجات على الصورة الأولى (الدرجة الكلية)، ومع ذلك فقد كان توزيع الدرجات متقاربًا؛ مما يدلّ على توازن عامّ بين الصورتين.

الشكل (29) مخطط الصندوق والمؤشر للدرجات الخام لصورتي الامتحان



وبشكل عام فإن هذه النتائج تشير إلى تقارب مستويات صورتي الامتحان مع وجود سهولة نسبية للصورة الأولى 2023/ 2024م، وتتفق هذه النتائج مع نتائج دراسات كلّ من: (الحبسي، 2018؛ الزعابية وحسن، 2018؛ Leoncio et al., 2023)، بينما اختلفت مع دراسة الشريفين (2009) في تقارب الدرجات الخام المتعادلة.

2. المعادلة بين صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام باستعمال درجات القدرة.

لإيجاد الدرجات المتعادلة باستعمال درجات القدرة (θ) فقد رُبِطَ بين الدرجات الحقيقية على الصورتين ودرجة القدرة المقابلة لها، ويوضح الجدول (24) درجة القدرة وما يقابلها من الدرجات الحقيقية على صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام.

الجدول (24) توزيع درجات القدرة وما يقابلها من درجات حقيقية على صورتي الامتحان

القدرة	الدرجة الحقيقية على الصورة الأولى	الدرجة الحقيقية على الصورة الثانية	القدرة	الدرجة الحقيقية على الصورة الأولى	الدرجة الحقيقية على الصورة الثانية
−4	5.91	4.56	0.1	57.99	56.11
−3.9	6.03	4.61	0.2	59.65	58.22
−3.8	6.16	4.67	0.3	61.15	60.13
−3.7	6.31	4.74	0.4	62.48	61.83
−3.6	6.48	4.83	0.5	63.65	63.32
−3.5	6.67	4.94	0.6	64.66	64.59
−3.4	6.90	5.07	0.7	65.54	65.66
−3.3	7.15	5.22	0.8	66.29	66.54
−3.2	7.43	5.41	0.9	66.93	67.27
−3.1	7.75	5.63	1	67.475	67.85
−3	8.11	5.89	1.1	67.94	68.31
−2.9	8.51	6.20	1.2	68.32	68.68
−2.8	8.97	6.56	1.3	68.64	68.97
−2.7	9.48	6.97	1.4	68.90	69.19
−2.6	10.05	7.44	1.5	69.12	69.37
−2.5	10.69	7.97	1.6	69.29	69.51

القدرة	الدرجة الحقيقية على الصورة الأولى	الدرجة الحقيقية على الصورة الثانية	القدرة	الدرجة الحقيقية على الصورة الأولى	الدرجة الحقيقية على الصورة الثانية
−2.4	11.40	8.58	1.7	69.43	69.62
−2.3	12.20	9.27	1.8	69.55	69.70
−2.2	13.09	10.04	1.9	69.64	69.77
−2.1	14.08	10.90	2	69.71	69.82
−2	15.17	11.86	2.1	69.77	69.86
−1.9	16.38	12.92	2.2	69.81	69.89
−1.8	17.72	14.10	2.3	69.85	69.91
−1.7	19.19	15.39	2.4	69.88	69.93
−1.6	20.78	16.81	2.5	69.90	69.94
−1.5	22.52	18.35	2.6	69.92	69.96
−1.4	24.38	20.03	2.7	69.94	69.97
−1.3	26.36	21.83	2.8	69.95	69.97
−1.2	28.46	23.78	2.9	69.96	69.98
−1.1	30.67	25.87	3	69.96	69.98
−1	32.96	28.09	3.1	69.97	69.99
−0.9	35.32	30.45	3.2	69.98	69.99
−0.8	37.74	32.92	3.3	69.98	69.99
−0.7	40.20	35.50	3.4	69.98	69.99
−0.6	42.66	38.15	3.5	69.99	69.99
−0.5	45.10	40.84	3.6	69.99	70
−0.4	47.49	43.54	3.7	69.99	70
−0.3	49.82	46.23	3.8	69.99	70
−0.2	52.05	48.86	3.9	69.99	70
−0.1	54.17	51.40	4	69.99	70
0	56.15	53.83	المتوسط	44.61	43.01
			الانحراف المعياري	26.07	27.30

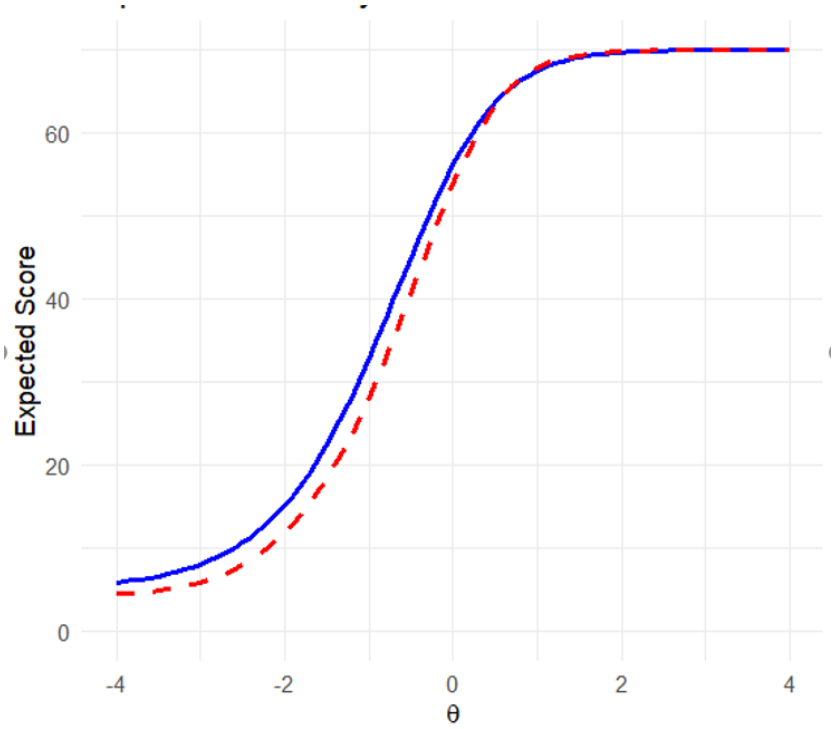
نلاحظ من الجدول (24) أن الإحصاءات الوصفية للدرجات الحقيقية على صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام متقاربة، حيث توزعت الدرجات الحقيقية على الصورة الأولى بمتوسط حسابي 44.61 وانحراف معياري 26.07، وتوزعت الدرجات الحقيقية على الصورة الثانية

بمتوسط حسابي 43.01 وانحراف معياري 27.30، وهذا يتفق مع ما أشار إليه Kolen and Brennan (2014) في أن توزيع الدرجات المعادلة للصورة الجديدة يتقارب مع توزيع الدرجات في الصورة المرجعية، وتتفق هذه النتائج مع نتائج دراسات كل من: (الشريفين، 2009؛ العبيدي، 2018؛ He et al., 2013).

ونلاحظ أيضًا من الجدول (24) أنه عند درجات القدرة المتدنية إلى المتوسطة تكون الدرجات الحقيقية على الصورة الأولى أعلى من الدرجات الحقيقية على الصورة الثانية، ونلاحظ عند درجات القدرة من المتوسطة إلى العليا تكون الدرجات الحقيقية على الصورة الثانية أعلى من الدرجات الحقيقية على الصورة الأولى، وهذا يدلّ على أن صورتي الامتحان لا تعملان بالطريقة نفسها عند جميع مستويات القدرة، بمعنى أن الصورة الأولى أفضل تمييزًا في المستويات الدنيا من القدرة، بينما تفوقت الصورة الثانية في المستويات العليا من القدرة، وهذا يشير إلى وجود ما يُسَمّى بالأداء التفاضلي للفقرات (DIF) Differential Item Functioning (Holland & Wainer, 1993). ويعود ذلك من وجهة نظر الباحث بسبب أن قيم التخمين بالنسبة للفقرات الموضوعية في الصورة الأولى كانت أعلى مما يتيح للأفراد ذوي القدرة المتدنية حلها بالتخمين بصورة كبيرة بينما يخفق فيها طلاب القدرات الأعلى، وأيضا عند النظر إلى منحنيات دالة معلومات فقرات صورتي الامتحان نجد أن دالة معلومات الكثير من فقرات الصورة الأولى منخفضة وهذا يؤدي إلى زيادة أخطاء القياس، فقد ينجح بالإجابة عن هذه الفقرات فرد ذو قدرة متدنية ويفشل في حلها فرد ذو قدرة أعلى.

ونلاحظ هذه الملاحظة أيضًا من الشكل (30)، حيث يوضح الشكل (30) درجات القدرة وما يقابلها من درجات حقيقية على صورتي الامتحان.

الشكل (30) درجات القدرة وما يقابلها من درجات حقيقية على صورتي الامتحان



وبشكل عام فإن كلّ هذه النتائج تشير إلى أن صورتي امتحان الرياضيات المتقدمة لدبلوم التعليم العام تعملان بشكل مختلف عند مستويات القدرة المختلفة، فمثلًا عند $\theta = -2$ تكون الدرجة الحقيقية على الصورتين على التوالي 15، و10 تقريبًا بفارق خمس درجات للصورة الأولى ، وعند $\theta = 0$ تكون الدرجة الحقيقية على الصورتين على التوالي 35، و32 تقريبًا بفارق ثلاث درجات للصورة الأولى، بينما $\theta = 2$ عند تكون الدرجة الحقيقية على الصورتين على التوالي 65، و67 تقريبًا بفارق درجتان للصورة الثانية، فعند مقارنة الدرجات الحقيقية على الصورتين عند مستويات مختلفة من القدرة (θ)، يتضح أن الصورة الأولى تمنح درجات أعلى عند القدرة المنخفضة والمتوسطة، في حين تبدأ الصورة الثانية في التفوق عند القدرة العالية. هذا التغير في اتجاه الفروق بين الدرجات الحقيقية يعزّز فرضية أن كلّ صورة تميّز شريحة معينة من مستويات القدرة بشكل أفضل، وهذا يتفق مع دراسة (الرحيل، 2013).

3. معادلة الدرجات الخام على الصورة الأولى لامتحان الرياضيات المتقدمة لدبلوم التعليم العام بما يقابلها من درجات قدرة ودرجات حقيقية على الصورة الثانية.

في هذه الحالة حُوِّلَت الدرجات الخام على الصورة الأولى بما يقابلها من درجات قدرة ودرجات حقيقية على الصورة الثانية، ويوضح الجدول (25) الدرجات الخام على الصورة الأولى وما يقابلها من درجات قدرة ودرجات حقيقية على الصورة الثانية.

الجدول (25) الدرجات الخام على الصورة الأولى وما يقابلها من درجات قدرة ودرجات حقيقية على الصورة الثانية

الدرجة الخام على الصورة الأولى	درجة القدرة	الدرجة الحقيقية على الصورة الثانية	الدرجة الخام على الصورة الأولى	درجة القدرة	الدرجة الحقيقية على الصورة الثانية
0	−3.187	5.463	36	−0.845	31.855
1	−3.078	5.715	37	−0.800	32.998
2	−3.051	5.787	38	−0.764	33.891
3	−3.037	5.823	39	−0.722	34.969
4	−2.846	6.428	40	−0.684	35.971
5	−2.805	6.580	41	−0.651	36.827
6	−2.638	7.301	42	−0.608	37.964
7	−2.509	7.975	43	−0.576	38.823
8	−2.453	8.303	44	−0.529	40.084
9	−2.401	8.632	45	−0.487	41.229
10	−2.263	9.603	46	−0.450	42.205
11	−2.179	10.274	47	−0.406	43.413
12	−2.060	11.341	48	−0.368	44.424
13	−2.034	11.588	49	−0.319	45.722
14	−1.951	12.445	50	−0.280	46.767
15	−1.904	12.959	51	−0.236	47.919
16	−1.832	13.793	52	−0.190	49.115
17	−1.764	14.639	53	−0.143	50.310
18	−1.685	15.686	54	−0.097	51.470
19	−1.645	16.236	55	−0.048	52.662

الدرجة الخام على الصورة الأولى	درجة القدرة	الدرجة الحقيقية على الصورة الثانية	الدرجة الخام على الصورة الأولى	درجة القدرة	الدرجة الحقيقية على الصورة الثانية
20	−1.616	16.660	56	0.002	53.851
21	−1.532	17.935	57	0.056	55.105
22	−1.480	18.760	58	0.110	56.292
23	−1.438	19.464	59	0.169	57.556
24	−1.364	20.751	60	0.233	58.836
25	−1.327	21.411	61	0.290	59.908
26	−1.273	22.430	62	0.359	61.115
27	−1.236	23.145	63	0.435	62.336
28	−1.207	23.720	64	0.511	63.431
29	−1.142	25.060	65	0.604	64.603
30	−1.082	26.345	66	0.716	65.775
31	−1.043	27.186	67	0.863	66.987
32	−1.011	27.918	68	1.039	68.018
33	−0.964	28.994	69	1.192	68.631
34	−0.919	30.061	70	1.666	69.575
35	−0.883	30.939		المتوسط	33.07
				الانحراف المعياري	19.90

نلاحظ من الجدول (25) أن الدرجات المتوقعة على الصورة الثانية توزعت بمتوسط حسابي 33.07 وانحراف معياري 19.90، ويوضح الجدول (26) كذلك معامل ارتباط بيرسون بين الدرجات المحولة، حيث إن معامل الارتباط (Correlation Coefficient) يدل على مدى قوة العلاقة واتجاهها بين نوعين من الدرجات.

الجدول (26) معامل ارتباط بيرسون بين الدرجات المحولة لصورتي الامتحان

المتغير	درجة القدرة	الدرجة الحقيقية على الصورة الثانية
الدرجة الخام على الصورة الأولى	0.99	0.99
الدرجة الحقيقية على الصورة الثانية	0.98	

نلاحظ من القيم في الجدول (26) أن قيم معاملات ارتباط بيرسون بين الدرجات المختلفة موجبة ومرتفعة، وهذا يدلّ على أن الدرجات الثلاث متسقة مع بعضها بدرجة قوية؛ لأن صورتي الامتحان تقيسان المحتوى نفسه (Anastasi & Urbina, 1997)، ونلاحظ بشكل عام اختلاف الدرجة الخام على الصورة الأولى المقابلة لدرجة قدرة معينة عن الدرجة الحقيقية على الصورة الثانية المقابلة لذات القدرة، وتتفق هذه النتائج مع دراستي كلّ من: (الشريفين، 2009؛ العبيدي، 2018) من حيث اختلاف الدرجات الخام عن الدرجات الحقيقية المعادِلة لها.

ملخص النتائج

تناولت هذه الدراسة معادلة درجات امتحان الرياضيات المتقدمة لدبلوم التعليم العام لعامي 2023/ 2024م و2024/ 2025م باستعمال تصميم المجموعات المتكافئة (العشوائية) باستعمال نظرية الاستجابة للفقرة، وذلك عن طريق الربط بين الدرجات الخام والدرجة الحقيقية ودرجات القدرة لصورتي الامتحان، وقد أظهرت النتائج من قيم التفلطح والالتواء أن بيانات صورتي الامتحان تتوزع توزيعًا طبيعيًا.

وأشارت النتائج أيضًا إلى وجود فروق ذات دلالة إحصائية بين الدرجات الخام لصورتي الامتحان، وأن صورتي امتحان الرياضيات المتقدمة لدبلوم العام تتمتّعان بمعدلات ثبات مرتفعة، ومن التحليل العاملي الاستكشافي والتحليل العاملي المتوازي أشارت النتائج إلى تحقيق صورتي الامتحان لافتراض أحادية البُعد لنظرية الاستجابة للفقرة، وقد تم التحقق من افتراض الاستقلال الموضعي لفقرات صورتي الامتحان باستعمال مصفوفة البواقي المعيارية.

وأكّدت النتائج مطابقة بيانات صورتي الامتحان للنموذج ثلاثي المعالم (3PLM) ونموذج التقدير الجزئي المعمم (G–PCM) لنظرية الاستجابة للفقرة، واتسمت خصائص فقرات صورتي الامتحان بصعوبة من سهلة إلى متوسطة وبتميز عالٍ، وقُدِّرَت قدرات مجموعتي الأفراد، حيث توزعت قدرات المجموعتين توزيعًا طبيعيًا، وذلك حسب قيم التفلطح والالتواء لقيم القدرة.

وأظهرت نتائج معادلة درجات صورتي امتحان الرياضيات المتقدمة سهولة نسبية لامتحان الصورة الأولى 2024/2023م، وأن فقرات الصورتين تعملان بشكل مختلف عند المستويات المختلفة للقدرة.

التوصيات والمقترحات

هدفت هذه الدراسة إلى معادلة درجات صورتين من امتحان مادة الرياضيات المتقدمة لدبلوم التعليم العام في سلطنة عُمان، وذلك باستعمال تصميم المجموعات العشوائية المتكافئة باستعمال نظرية الاستجابة للفقرة، ومن النتائج التي توصلت إليها هذه الدراسة يمكن الخروج بالتوصيات والمقترحات الآتية:

أولًا. التوصيات:

- استفادة المختصين في وزارة التربية والتعليم من دراسات معادلة درجات الامتحانات؛ لإجراء المقارنات الموضوعية بين مستويات الطلبة عبر الأعوام الدراسية المختلفة.
- استفادة المختصين في وزارة التربية والتعليم من دراسات معادلة درجات الامتحانات في تطوير الفقرات الامتحانية لمادة الرياضيات المتقدمة، وتحسين خصائصها؛ مما يؤدّي إلى تحسين جودة القياس.
- استفادة المختصين في وزارة التربية والتعليم من دراسات معادلة درجات الامتحانات في تحسين خصائص الصور الاختبارية لمادة الرياضيات المتقدمة من خلال توزيع الفقرات على كامل متصل السمة.
- استفادة وزارة التربية والتعليم من دراسات معادلة درجات الامتحانات لبناء بنوك أسئلة تشتمل على فقرات محددة المعالم تُوظَّف لأغراض القياس المختلفة.

ثانيًا. المقترحات:

- إجراء دراسات معادلة باستعمال طرائق أخرى من طرائق المعادلة وعدم الاعتماد على طريقة واحدة؛ لأن طرائق المعادلة تتكامل فيما بينها.
- إجراء دراسات معادلة درجات صور امتحانية لمواد دراسية أخرى ولا سيما المواد الدراسية التي تُعَدّ ركيزة أساسية لركائز تقنيات الثورة الصناعية الرابعة (مواد العلوم والرياضيات وتقنية المعلومات).
- إجراء دراسات معادلة لصور امتحانية على مستوى كلّ محافظة تعليمية؛ وذلك للمقارنة بينها وبين المعادلة العامة.
- إجراء دراسات الكشف عن الفقرات ذات الأداء التفاضلي في امتحانات دبلوم التعليم العام وأثر هذه الفقرات على دقة المعادلة.
- اجراء دراسات معادلة للاختبارات والمسوحات الوطنية والدولية المتوفرة لدى الجهات المختصة.

قائمة المراجع

أولًا. المراجع العربية:

بوصلب، عبد الحكيم (2023). قراءة في مدلول العلامة المدرسية: المدرسة الجزائرية نموذجًا. *مجلة طبة للدراسات العلمية الأكاديمية. المركز الجامعي سي الحواس بريكة. 6*(2)،277-302.

التقي، أحمد محمد (2008). *النظرية الحديثة في القياس*. دار المسيرة للنشر والتوزيع والطباعة.

جاد الله، أشرف سعد، وعلي، محمد سعد، وعبد الواحد، إبراهيم سيد (2021). أثر اختلاف طريقة المعادلة في معادلة صورتي اختبار تحصيلي وفق نظرية الاستجابة للمفردة. *مجلة التربية. كلية التربية. جامعة الأزهر2*(192)، 600-640.

الحبسي، نبيل بن سيف (2018). *معادلة درجات امتحان دبلوم التعليم العام لمادة الفيزياء في سلطنة عُمان باستخدام نظرية الاستجابة للمفردة* [رسالة ماجستير غير منشورة]. جامعة السلطان قابوس، سلطنة عُمان.

الحجوري، عيسى بن سلامة، وعبد الوهاب، محمد محمود (2024). أثر نموذج نظرية الاستجابة للمفردة في دقة معادلة درجات الاختبارات، *مجلة العلوم التربوية والنفسية، 8*(9)، 98- 115.

الحربي، عيسى بن جود الله، وعليان، خليل (2009). *أثر دقة تمثيل الفقرات للمحتوى ونسبة الفقرات المشتركة وطرق توزين الفقرات على دقة معادلة درجات الصور للاختبار عندما يكون التصميم المستخدم هو الجذع المشترك (الفقرات المشتركة)* [أطروحة دكتوراة منشورة، الجامعة الأردنية]. قاعدة معلومات دار المنظومة.

حسن، علي صلاح، وحسن، سميرة محمد، وعيسى، عماد أحمد (2016). *أثر أداء الفقرات التفاضلي للنوع والعمر في دقة معادلة الاختبار في ضوء نموذجي راش وفيشر للاستجابة للمفردة* [أطروحة دكتوراه منشورة، جامعة أسيوط]. قاعدة معلومات دار المنظومة.

حمادنة، إياد محمد (2012). أثر تجانس المحتوى وطول الجذع المشترك على دقة معادلة اختبار ثنائي التدريج ومتعدد المستوى في مبحث الرياضيات. *مجلة المنارة للبحوث والدراسات، 18*(4)، 9-43.

حمادنة، راتب صايل (2013). أثر وجود أداء تفاضلي في الفقرات المرساوية على دقة المعادلة العمودية لاختبار أوتيس لينون للقدرة العقلية. *دار سمات للدراسات والأبحاث*، 2(8)، 754 – 771.

الدوسري، راشد حماد (2001). معادلة الاختبارات، مفهومها، وطرقها، ومشكلات تطبيقها. *مجلة العلوم النفسية والتربوية، 2*(4)، 106- 141.

الرحيل، راتب صايل (2013). أثر وجود اداء تفاضلي في الفقرات المرساوية على دقة المعادلة العمودية لاختبار أوتيس لينون للقدرة العقلية. *المجلة الدولية التربوية المتخصصة، 2*(8)، 754- 771.

الرواجفة، عبد الله علي، والطراونة، صبري حسن خليل (2016). *المعادلة العمودية لاختبار متعدد المستويات للصفوف (العاشر والأول الثانوي العلمي والثاني الثانوي العلمي) باستخدام طريق النظرية الحديثة في القياس* [رسالة ماجستير منشورة، جامعة مؤتة]. قاعدة معلومات دار المنظومة.

الزعابية، أمل بنت خميس، وحسن، عبد الحميد (2018). *معادلة كتيبات الدراسة الدولية لقياس مهارات القراءة PIRLS للصف الرابع الأساسي في سلطنة عُمان باستخدام نظرية الاستجابة للمفردة* [رسالة ماجستير منشورة، جامعة السلطان قابوس] قاعدة معلومات دار المنظومة.

الزعابية، أمل بنت خميس، والمحرزي، راشد بن سيف، وحسن، عبد الحميد (2020). معادلة كتيبات اختبار الدراسة الدولية لقياس مهارات القراءة (PIRLS2011) سلطنة عُمان باستخدام نظرية الاستجابة للمفردة. *مجلة العلوم التربوية*، (15)، 68 -93.

الشريفين، نضال كمال (2009). معادلة درجات نماذج مختلفة من اختبار الكفاءة اللغوية في اللغة الإنجليزية لدى طلبة جامعة اليرموك. *مجلة جامعة أم القرى للعلوم التربوية والنفسية، 1*(2)، 11 – 62.

الصمادي، إسماعيل محمد، وعدس، عبد الرحمن (2008). *أثر طريقة اختيار الفقرات في اختبار الجذع المشترك على دقة معادلة اختبار متعدد المستوى في الرياضيات للمرحلة الأساسية في الأردن* [أطروحة دكتوراة منشورة، جامعة عمان العربية]. قاعدة معلومات دار المنظومة.

عبد الوهاب، محمد محمود (2015). أثر اختلاف معامل ثبات صورتي اختبار في دقة معادلة درجاتهما باستخدام الطرق القائمة على نظرية الاستجابة للمفردة. *مجلة كلية التربية، 25*(3)، 23 – 54.

العبيدي، شيماء صلاح (2018). معادلة اختبارات الذكاء المتقدمة باستعمال الدرجة الحقيقية في النظرية الكلاسيكية. *حوليات آداب عين شمس، 46*، 140 – 171.

علام، صلاح الدين (2005). *نماذج الاستجابة للمفردة الامتحانية أحادية البُعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي*. دار الفكر العربي.

العنزي، خالد بن طخطيخ (2014). أثر حجم العينة في دقة معادلة درجات صورتي اختبار ثنائي التدريج وفق تصميم الجذع المشترك. *عالم التربية، 15*(47)، 235 – 264.

العنزي، خالد بن طخطيخ (2015). مقارنة طريقة تكر الخطية وطريقة الدرجات الحقيقية في دقة معادلة درجات صورتي اختبار ثنائي التدريج. *المجلة العربية للعلوم الاجتماعية، 3*(7)، 215– 244.

العنزي، خالد بن طخطيخ (2018). أثر طول الاختبار في دقة معادلة صورتي اختبار في الرياضيات باستخدام نظرية الاستجابة للمفردة. *المجلة العربية للعلوم الاجتماعية، 1*(13)، 19 – 55.

العنزي، سامي شطيط (2021). دراسة مقارنة لتصميمات المعادلة في عملية معادلة درجات الاختبارات متعددة الأبعاد في ظل جائحة كرونا باستخدام نظرية الاستجابة للمفردة. *مجلة القراءة والمعرفة*، (233)، 239 – 328.

عودة، أحمد سليماني (2010). *القياس والتقويم في العملية التدريسية* (ط.4). دار الأمل للنشر والتوزيع.

القرني، أحمد سالم (2023). أثر طريقة معالجة القيم المفقودة وتقدير القدرة للأفراد على دقة معادلة الاختبارات في نظرية الاستجابة للمفردة. *العلوم التربوية، 3*(4)، 623 – 686.

كمال، أحمد عبد البديع (2022). حجم التأثير والفاعلية في البحوث التجريبية. *المجلة الدولية لبحوث الإعلام والاتصالات، 2*(3)، 3-27.

المحرزي، راشد بن سيف (2014). المعادلة بين نتائج طرق المعادلة الكلاسيكية لدرجات نماذج اختبارات القدرات العامة باستخدام تصميم الجماعات المتكافئة. *رسالة الخليج العربي، 35*(134)، 15-42.

المحرزي، راشد بن سيف (2015). المفاضلة بين الدرجات المكافئة لنماذج اختبار القدرة العامة باستخدام طرق المعادلة الكلاسيكية في تصميم المفردات المشتركة بجماعات غير متكافئة. *مجلة العلوم التربوية والنفسية، 16*(3)، 393-429.

المحروق، يوسف عبد العاطي (2016). أثر حجم العينة وطول الاختبار في دقة معادلة درجات الاختبارات متعددة الحدود باستخدام طريقة كيرنيل. *مجلة العلوم التربوية والنفسية، 17*(3)، 201-228.

المحروق، يوسف عبد العاطي (2017). دقة معادلة درجات الاختبارات متعددة الاستجابة في ضوء اختلاف حجم العينة ومستوى الصعوبة. *العلوم التربوية، 44*(3)، 223 – 237.

المحروق، يوسف عبد العاطي، وعليان، خليل محمد (2011). *مقارنة طريقة كيرنل والمئينات وطرق نظرية استجابة الفقرة عند استخدام تصميم الفقرات المشتركة في دقة معادلة درجات الاختبارات متعددة الحدود* [أطروحة دكتوراة منشورة، الجامعة الأردنية]. قاعدة معلومات دار المنظومة.

المدانات، رائد فايز، والكيلاني، عبد الله زيد (2008). *أثر طريقة المعادلة باستخدام الجذع المشترك وعدد فقراته وحجم العينة على القيم المعادلة والخطأ في المعادلة بين صورتي اختبار في الفيزياء* [أطروحة دكتوراه منشورة، جامعة عمان العربية]. قاعدة معلومات دار المنظومة.

وزارة التربية والتعليم (2022أ). *نشرة توجيهية حول تدريس مناهج مواد الرياضيات في ضوء تطبيق سلاسل كامبريدج العالمية*. سلطنة عُمان.

وزارة التربية والتعليم (2022ب). *نشرة توضيحية بشأن تطبيق مناهج كامبردج لمادتي العلوم والرياضيات*. سلطنة عُمان

وزارة التربية والتعليم (2024أ). *وثيقة تقويم تعلم الطلبة في مواد الرياضيات للصفوف 11و12 العام 2024-2025*. سلطنة عُمان.

وزارة التربية والتعليم (2025). *الكتاب السنوي للإحصاءات التعليمية- الإصدار 55*. سلطنة عُمان.

وزارة التربية والتعليم(2024ب). *وثيقة تقويم تعلم الطلبة في مادة الرياضيات (الأساسية/ المتقدمة) للصفوف 11-12*. سلطنة عُمان.

الوليلي، إسماعيل (2005). تكافؤ درجات اختبارات في ضوء نظريتي القياس الكلاسيكية والحديثة (دراسة سيكومترية مقارنة). *مجلة كلية التربية، جامعة بنها، 15*(63)، 99–149.

ثانيًا. المراجع الأجنبية:

Anastasi, A., & Urbina, S. (1997). *Psychological testing (7th ed.).* Pearson Education, Prentice Hall.

Baker, F. (2001). *The Basics of Item Response Theory. (2rd ed*). ERIC Clearinghouse on Assessment and Evaluation. USA.

Baker, F. B., & Kim, S.-H. (2004). *Item Response Theory: Parameter Estimation Techniques (2nd ed.)*. CRC Press.

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation, 18*(6), 1–13.

Bazirhan, U. & von Davier, M., Fishbein, B. (2023). *TIMSS 2023 Technical Report – Methods & Procedures (Chapter 11)*. TIMSS & PIRLS International Study Center, Boston College.

Bond, T., & Fox, C. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Erlbaum.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Rinehart and Winston Inc.

De Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. The Guilford press.

Dorans, N. J., Moses, T. P., & Eignor, Daniel R. (2010). *Principles and Practices of Test Score Equating*. Educational Testing Service.

Dorans, N. J., Pommerich, M., & Holland, P. W. (2007). *Linking and Aligning Scores and Scales*. Springer Science & Business Media.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399–412.

Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for Psychologists Multivariate Applications Book Series.* Lawrence Erlbaum Associates, Inc.

Felan, G. D. (2002, February14-16). *Test equating: Mean, linear، Equipercentile, and Item response theory*. The annual meeting of the southwest educational research association. Austin, TX.

Fellinghauer, C., Debelak, R., & Strobl, C. (2023). What Affects the Quality of Score Transformations? Potential Issues in True-Score Equating Using the Partial Credit Model. *Educational and Psychological Measurement, 83*(6), 1249 – 1290.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE Publications

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update (10th ed)*. Boston: Pearson

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and Applications*. Springer Science & Business Media.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Publications.

Harlow, L. L. (2005). *The essence of multivariate thinking: Basic themes and methods*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hayes, A. F. (2012*). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7*(2), 191–205

He, Y., Cui, Z., Fang, Y., & Chen, H. (2013). *Applied Psychological Measurement, 37*(7), 522 – 540.

Holland, P. W, Dorans, N. J., & Petersen, N. S. (2007). Equating Test Scores. *Handbook of Statistics,* (*26*), 169 – 203.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning. Hillsdale*, NJ: Lawrence Erlbaum Associates.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2),179–185.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Istiyono, E., Dwandaru, W., Setiawan, R., & Megawati, I. (2020). Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use. *European Journal of Educational Research, 9*(1),91-101.

Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd ed.). Springer Series in Statistics

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36.

Kolen, M. J, & Brennan, R. J. (2014). *Test Equating: Methods & Practices* (3rd Ed.). NY: Springer-Verlag.

Leoncio, W., Wiberg, M., & Battauz, M. (2023). Evaluating Equating Transformations in IRT Observed-Score and Kernel Equating Methods. *Applied Psychological Measurement, 47*(2), 123 – 140.

Linacre, J. M. (2002). What do Infit and Outfit, Mean-square and Standardized mean?. *Rasch Measurement Transactions, 16*(2), 878.

Linacre, J. M. (2017). *Winsteps Rasch Measurement Software: Misfit Diagnosis.* Retrieved from https://www.winsteps.com/winman/misfitdiagnosis.htm.

Linacre, J. M. (2024). *A user's guide to WINSTEPS MINISTEP Rasch-model computer programs (Version 5.5.5)*. Winsteps.

Immekus, J. C., Snyder, K. E., & Ralston, P. A. (2019). Multidimensional item response theory for factor structure assessment in educational psychology research. *Frontiers in Education, 4*(45).

Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hillsdale, N. J. Erlbaum.

Martin, A. J., & Lazendic, G. (2018). Computer-adaptive testing: Implications for students' achievement, motivation, engagement, and subjective test experience. *Journal of Educational Psychology, 110*(1), 27–45.

Ogunsakin, I. B. & Shogbesan, Y. (2018). Item Response Theory (IRT): A Modern Statistical Theory for Solving Measurement Problem in 21st Century. *International Journal of Scientific Research in Education, 11*(3B), 627-635.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. SAGE Publications.

Reckase, M. D. (2009*). Multidimensional Item Response Theory*. Springer Science & Business Media.

Schumacker, R. E., & Lomax, R. G. (2016). *A beginner's guide to structural equation modeling (4th ed.)*. Routledge Taylor and Francis Group.

Song, T. (2009). Investigating different item response models in equating the Examination for the Certificate of Proficiency in English (ECPE). *Spaan Fellow Working Papers in Second or Foreign Language Assessment, 7*(1), 85–98

Sun, T. & Kim, S.Y. (2024). Evaluating Equating Methods for Varying Levels of Form Difference. *Educational and Psychological Measurement, 84*(3), 510- 529.

Von Davier, A. A. (2011). *Statistical Models for Test Equating, Scaling, and Linking*. Springer Science & Business Media.

Wallin, G., & Wiberg, M. (2025). Smoothing of Bivariate Test Score Distribution: Model Selection Targeting Test Score Equating. *Journal of Educational and Behavioral Statistics, 0*(0), 1-35.

Wallmark, J., Josefsson, M. & Wiberg, M. (2023). Efficiency Analysis of Item Response Theory Kernel Equating for Mixed-Format Tests. *Applied Psychological Measurement, 47*(8), 496 – 512.

Wiberg, M. (2016). Alternative Linear Item Response Theory Observed-Score Equating Methods. *Applied Psychological Measurement, 40*(3), 180 – 199.

Wilson, M. (2005*). Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Xu, Y. (2009). *Measuring change in jurisdiction achievement over time: Equating issues in current international assessment programs* [unpublished doctoral dissertation]. University of Toronto.

Yamamoto, K., & Kulick, E. (2000). *TIMSS 1999 Technical Report*. Chestnut Hill, MA: International Study Center, Boston College.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement, 8*(2), 125–145.

Zhonghua, Z. (2010). *Comparison of different equating methods and an application to link testlet-based tests* [unpublished doctoral dissertation]. The Chinese University of Hong Kong.

الملاحق

ملحق (1) رسالة طلب تسهيل مهمة باحث من جامعة الشرقية

جامعة الشرقية
ASHARQIYAH UNIVERSITY

كلية الاداب والعلوم الإنسانية

التاريخ: 2025/02/12

إلى من يهمه الأمر

تحية طيبة... وبعد

الموضوع/ تسهيل مهمة باحث

يرجى التكرم بتسهيل مهمة الطالب **علي بن عامر بن سلّوم الهنائي** المسجل في برنامج ماجستير في التربية: تخصص القياس والتقويم بجامعة الشرقية، كلية الأداب والعلوم الإنسانية، بقسم علم النفس من أجل الحصول على البيانات اللازمة لدراسة وصفية بعنوان:"**معادلة درجات امتحان دبلوم التعليم العام لمادة الرياضيات المتقدمة في سلطنة عُمان باستخدام نظرية الاستجابة للفقرة**"، إذ تتطلب الدراسة الحصول على الدرجات الفرعية لطلبة دبلوم التعليم العام في مادة الرياضيات المتقدمة/ الفصل الأول خلال العام الأكاديمي الحالي والعام الأكاديمي الماضي. علماً بأن البيانات تستخدم لأغراض البحث العلمي فقط.

شاكرين ومقدرين تعاونكم الدائم.

وتفضلوا بقبول فائق الاحترام والتقدير

د. محمد بن خلفان الصقري

عميد كلية الآداب والعلوم الانسانية

ملحق (2) رسالة طلب بيانات من وزارة التربية والتعليم

الفاضل / مدير دائرة الدراسات التربوية والتعاون الدولي المحترم

السلام عليكم ورحمة الله وبركاته...

الموضوع / طلب تسهيل مهمة باحث لإجراء دراسة بحثية لاستكمال متطلبات دراسة الماجستير

بالإشارة إلى الموضوع اعلاه ، وبناء على تواصلي مع المعنيين بدائرة الدراسات التربوية والتعاون الدولي ، وكذلك تواصلي المسبق مع المختصين في المركز الوطني للقياس والتقويم بوزارة التربية والتعليم . أتقدم إليكم بطلب تسهيل مهمة إنجاز الدراسة البحثية لاستكمال دراسة الماجستير بعنوان **(معادلة درجات امتحان الدبلوم العام لمادة الرياضيات المتقدمة باستخدام نظرية الاستجابة للفقرة في سلطنة عُمان) .** حيث تتطلب الدراسة الحصول على البيانات والمعلومات التالية :

- **درجات الدبلوم العام لمادة الرياضيات المتقدمة للفصل الدراسي الأول من العام الدراسي 2023 / 2024م**
- **درجات الدبلوم العام لمادة الرياضيات المتقدمة للفصل الدراسي الأول من العام الدراسي 2024 / 2025م**
- **امتحان الدبلوم العام لمادة الرياضيات المتقدمة للفصل الدراسي الأول مع نموذج الإجابة للعام الدراسي 2023 / 2024م**
- **امتحان الدبلوم العام لمادة الرياضيات المتقدمة للفصل الدراسي الأول مع نموذج الإجابة للعام الدراسي 2024 / 2025م**
- **كشف بإحصائيات أعداد الطلاب المتقدمين لامتحان الدبلوم العام لمادة الرياضيات المتقدمة للفصل الدراسي الأول للعامين الدراسيين 2023 / 2024م و2024 / 2025م .**
- **المنشورات والادلة المتعلقة بإعداد المفردات الامتحانية لمناهج كامبريدج لدبلوم التعليم العام.**

حيث أن المختصين المركز الوطني للقياس والتقويم أبدوا موافقتهم المبدئية بإتاحة المعلومات والبيانات اللازمة لإنجاز الدراسة وذلك من خلال تواصلي الشخصي المسبق معهم.

لذا ألتمس منكم مساعدتي في انجاز هذه الدراسة البحثية من خلال الموافقة على طلبي هذا .

وتفضلوا بقبول فائق الاحترام والتقدير

مقدم الطلب: علي بن عامر بن سلّوم الهنائي

طالب ماجستير تخصص القياس والتقويم جامعة الشرقية

مكان العمل : وزارة التربية والتعليم (أخصائي تقييم برامج إثراء محلي)

تم بحمد الله